MASTERING ECONOMICS

JACK HARVEY

Mastering

Economics

Macmillan Master Series

Accounting
Arabic
Astronomy
Background to Business
Banking
Basic Management
Biology
British Politics
Business Communication
Business Law
Business Microcomputing
C Programming
Catering Science
Catering Theory
Chemistry
COBOL Programming
Commerce
Computer Programming
Computers
Economic and Social History
Economics
Electrical Engineering
Electronics
English as a Foreign Language
English Grammar
English Language
English Literature
French 1
French 2

German 1
German 2
Hairdressing
Human Biology
Italian 1
Italian 2
Japanese
Manufacturing
Marketing
Mathematics
Mathematics for Electrical and Electronic Engineering
Modern British History
Modern European History
Modern World History
Pascal Programming
Philosophy
Physics
Psychology
Restaurant Service
Science
Secretarial Procedures
Social Welfare
Sociology
Spanish 1
Spanish 2
Spreadsheets
Statistics
Study Skills
Word Processing
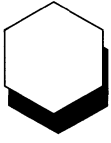
Mastering

Economics

Fourth Edition

Jack Harvey

M
MACMILLAN

⬡ Contents

List of Tables

⬡ List of Figures

⬡ Preface to the First Edition

Mastering Economics is written for those who are embarking on a study of economics. While its main aim is to cover basic examinations in schools and colleges, it will also meet the requirements of most professional syllabuses and will provide preliminary reading for university students.

For the general reader seeking guidance in understanding economic problems and policies, the book sets out the essential background and introduces in simple form the techniques used by economists.

Particular attention has been paid to current views regarding the defects of the 'laissez-faire' system in solving the economic problem, the role the government can play in a 'mixed' economy, and the increase in the importance of the public sector over the past thirty-five years – subjects which are inclined to be glossed over in elementary texts.

I would like to thank Andrew Leake, senior economics master at Latymer Upper School, and David Whitehead, lecturer in education at the University of London, who read the original typescript and made many helpful suggestions.

⬡ Preface to the Fourth Edition

Tables and facts have been brought up to date and the text has been revised to cover such changes as have taken place in the working of institutions and the evolution of economic policy in recent years.

The opportunity has also been taken to rationalise the role of the government in a market economy, especially as regards improving the allocation of resources and in formulating policy to achieve both full employment and price stability.

A new section on the environment surveys the economic aspects of external costs and benefits, the role of cost-benefit analysis, and suggests economic approaches to the problems of conservation and pollution.

While the analysis of the factors which determine aggregate demand have been retained, the failure of demand management to combat inflation has necessi-tated a new theoretical approach. Hence a simplified model relates aggregate demand and aggregate supply to explain why prices may continue to rise even though aggregate demand is falling, and introduces the reasoning behind supply-side economics.

The European Community is also discussed in more detail. An additional final chapter surveys briefly the major economic problems facing the UK government today.

JACK HARVEY

Acknowledgement
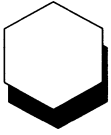
Part I
Introduction

① Solving the problem of scarcity

1.1 The economic problem

(a) Wants and limited resources

Have you been window-shopping lately, gazing longingly at the various goods on display? If only you had the means to buy them! This is the *economic problem* – unlimited wants, very limited means. You just can't get a quart out of a pint pot.

While we can never completely overcome the difficulty, we can, by 'economising', make the most of what we have. Thus the housewife buys that assortment of goods which will give a maximum satisfaction from her limited housekeeping allowance. Similarly, the student strives to make his grant go as far as possible. And the businessman takes decisions which will achieve the maximum return on capital. The government, too, has to plan its spending in order to make the most of the funds at its disposal.

(b) Opportunity cost

Thus economics is concerned with the problem of choice – the decisions forced upon us by the smallness of our resources compared with our wants. But choice involves sacrifice. If the newspaper boy spends his earnings on a football, he will have to postpone buying the table-tennis bat he also wants. The schoolgirl who works in a store on a Saturday has to forgo the game of tennis she would otherwise have played. When the farmer sows a field with wheat, he accepts that he will have to go without the barley it could have grown.

Because resources are limited, having 'this' means going without 'that', or, as the Yorkshireman says, 'There's no owt for nowt in this world'. We speak, therefore, of *opportunity cost* – the cost of something in terms of the best alternative gone without.

Usually economising does not mean a complete rejection of one good in favour of another, but rather deciding to have a little bit more of one and a little less of the other. In short, as we shall see it involves choice at the *margin*.

(c) 'Free' and 'scarce' goods

Few goods are so plentiful that nobody will give anything for them. Air is one of the few exceptions. In some years, too, there is such an abundant apple harvest that a farmer says, 'Help yourself'. Such goods are termed 'free' goods. But most goods are 'scarce – they can be obtained only by going without something else. With these goods we have to economise, so they are referred to as 'economic goods'. They are the subject-matter of economics – the study of how people allocate their limited resources to provide for their wants. It is against this backcloth
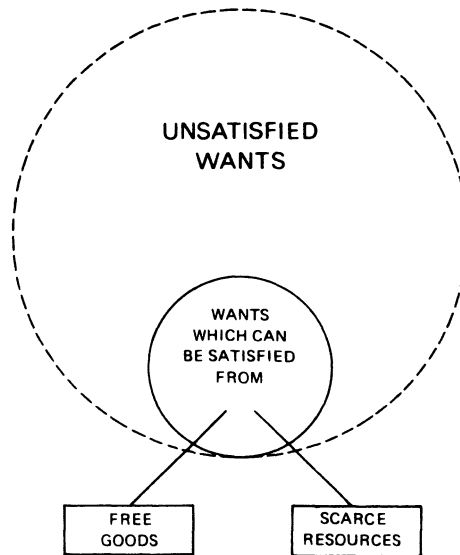
Figure 1.1   The economic problem

of limited resources that all economic decisions by consumers and firms have to
be made.

1.2   Economic systems

(a) The role of the economic system
In primitive economies, the individual uses his resources directly to provide what
he wants. Thus Robinson Crusoe had to decide how much time to spend hunting,
fishing, growing corn and relaxing in the sun according to the strength of his
preferences for meat, fish, bread and leisure. Similarly, in a subsistence economy
the farmer's output is mainly for his own family's needs.

Today, however, decisions as to what shall be produced are linked only indi-
rectly with the actual consumer. Man now specialises in production, obtaining
the variety of goods he wants by exchange. Thus, on the one hand, we have what
we will call 'households', the units which both consume goods and services and
supply the resources, such as labour, to produce them. On the other hand, we
have 'firms', the organisations which decide what goods and services to produce,
and use accordingly the resources supplied by households (Figure 1.2).

But if the greatest possible satisfaction is to be obtained from limited resources
there must be a link between households and firms. Put briefly, the following
questions have to be answered:
   (i) *What* goods and services shall firms produce?
  (ii) *How much* of each good and service shall be produced?
 (iii) *How* shall the goods and services be produced?
  (iv) How shall products be *divided* between households?
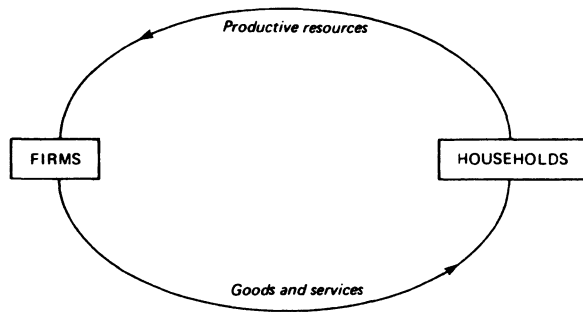
Figure 1.2 The flow of goods and resources in an economic system

To solve these problems we need some form of economic system; in short, the economic system provides the link between households and firms.

(b) Different forms of economic system

Man's first exchanges were quite simple: there was a direct swap of one good for another – a 'market' was established (see p. 29). Eventually a 'go-between' – money – was developed, allowing goods to be 'priced' and sold more easily. The subsistence economy had now evolved into the *market economy*, where answers to the above questions follow from people's decisions in the market.

In contrast to the market economy, there is the *command* or *centrally directed economy*, where the state decides what to produce and directs the factors of production accordingly. Furthermore, what is produced is distributed according to the decisions of the central body, the emphasis being 'to each according to his need' rather than on financial ability to pay.

Our task now is to examine in turn the respective strengths of these two systems.

1.3 The market economy

(a) Outline of the market mechanism

In the market economy, emphasis is laid on the freedom of the individual, both as a consumer and as the owner of resources.

As a consumer he expresses his choice of goods through the price he is willing to pay for them. As the owner of resources used in production (usually his own labour) he seeks to obtain as large a reward as possible. If consumers want more of the good than is being supplied at the current price, this is indicated by their 'bidding-up' the price. This increases the profits of firms and the earnings of factors producing that good. As a result, resources are attracted into the industry, and supply expands. On the other hand, if consumers do not want a particular good, its price falls, producers make a loss, and resources leave the industry.

The *price system* therefore indicates the wishes of consumers and allocates the community's productive resources accordingly (Figure 1.3). There is no direction of labour; people are free to work wherever they choose.
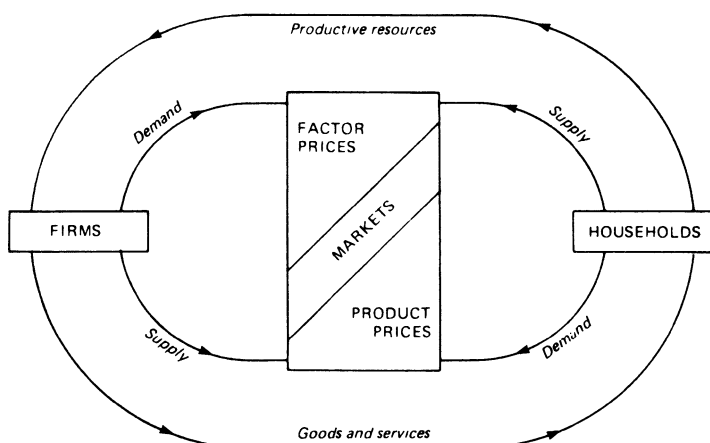
Figure 1.3   The allocation of products and resources through the market economy

Efficiency is achieved through the profit motive: owners of factors of production sell them at the highest possible price, while firms keep production costs as low as they can in order to obtain the highest profit margin.

Factor earnings decide who is to receive the goods produced. If firms produce better goods or improve efficiency, or if workers make a greater effort, they receive a high reward, giving them more spending power to obtain goods in the market.

In this way the price system acts, as it were, like a marvellous computer, registering people's preferences for different goods, transmitting those preferences to firms, moving resources to produce the goods, and deciding who shall obtain the final products. Thus, through the motivation of private enterprise, the four problems inherent in economising are solved automatically.

(b) Defects of the market economy
In practice the market economy does not work quite so smoothly as this. Nor are its results entirely satisfactory. We speak of 'market failure'.

First, some vital *community goods*, such as defence, police, justice and national parks, cannot be adequately provided through the market. This is mainly because it would be impossible to charge a price since 'free-riders' cannot be excluded. Indeed, in most advanced countries, the state usually goes further. Thus it may take responsibility for *public goods*, such as TV programmes and parks, where there is no reduction in the quantity available for others when one person has more, and the cost can be covered by taxation. Moreover, it usually provides a safety-net when people are unemployed, sick or old, and gives assistance towards *merit goods*, such as education, housing, museums and libraries, on which people might otherwise underspend.

Second, the consumers with the most money have the greatest pull in the market. As a result, resources may be devoted to producing luxuries for the rich to

the exclusion of necessities for the poor. While this is really brought about by the unequal distribution of wealth and income rather than by the market system, the fact is that the latter tends to produce, and even to increase, such inequality.

Third, the *competition* upon which the efficiency of the market economy depends *may break down*. An employer may be the only buyer of a certain type of labour in a locality. If so, he is in a strong position when negotiating rates of pay with individual workers. The state may therefore have to intervene, e.g. with minimum-wage requirements, as in agriculture. Similarly, on the selling side, one seller may be able to exclude competitors. This puts the consumer in a weak position because he cannot take his custom elsewhere.

Fourth, *competition itself may sometimes lead to inefficiency.* Small units may not be able to secure the advantages of large-scale production. Duplication of research and competitive advertising may waste resources. Uncertainty as to rivals' plans may hold back investment.

Fifth, *consumers' sovereignty may be distorted* by large firms which use extensive advertising simply to persuade people that their goods are just what they want.

Sixth, in practice the price mechanism may function sluggishly, through *imperfect knowledge* or *immobility of factors of production* (see p. 136). As a result supply is slow to respond to changes in demand.

Seventh, the private-profit motive does not always ensure that *public* well-being (as distinct from the sum total of *private* wealth) will be maximised. There may be 'spillover' benefits (often referred to as *externalities*). Thus, in providing a car park, a supermarket attracts customers; but there is an additional benefit in that it reduces congestion for all road users. On the other hand there may be *external costs*. A manufacturer does not consider the soot which falls from his factory chimney onto nearby washing-lines; but, although not a cost to him, it is one to the surrounding community.

Lastly, in a market economy where individuals decide what to produce, resources may remain *unemployed* because firms as a whole consider that profit prospects are poor.

1.4 The command economy

(a) Central decision-making
With the command economy, the decisions regarding what? how much? how? and for whom? are taken by an all-powerful planning authority. It estimates the assortment of goods which it considers people want and directs resources into producing them. It also decides how the goods produced shall be distributed among the community. Thus economic efficiency largely depends upon how accurately wants are estimated and resources allocated.

(b) Merits of the command economy
The merits of the command economy correspond closely to the defects of the market economy. The central planning authority can: (i) allow for the uneven distribution of wealth when planning what to produce and in rewarding the
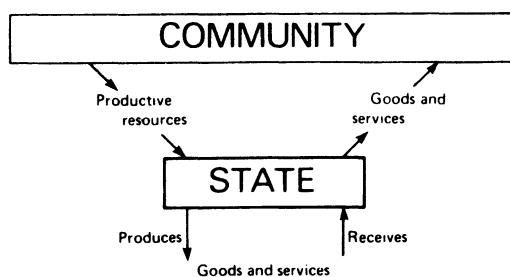
Figure 1.4   The allocation of resources and products through a command economy

producers; (ii) ensure that adequate resources are devoted to community, public and merit goods; (iii) eliminate the inefficiencies resulting from competition; (iv) use its monopoly powers in the interests of the community, e.g. by securing the advantages of large-scale production, rather than restricting output to maximise profits; (v) use advertising to inform rather than simply to persuade or 'brain-wash'; (vi) allow for external costs and benefits when deciding what and how much to produce; and (vii) employ workers in order to keep them occupied although to do so may be unprofitable in the narrow sense.

(c) Defects of the command economy
Nevertheless the command economy has inherent defects which lay it open to criticism on both economic and political grounds.

First, estimating the satisfaction derived by individuals from consuming different goods is impossible – although some help can be obtained by introducing a pricing system through markets, changes in prices signalling changes in wants.

Second, many officials are required to estimate wants and to direct factors of production. Inasmuch as such officials are not needed in a market economy they represent wasted factors of production. Moreover, the use of officials may lead to bureaucracy – excessive form-filling, 'red tape', slowness in coming to decisions and an impersonal approach to consumers. At times, too, officialdom has been accompanied by corruption.

Third, even when wants have been decided upon, difficulties of co-ordination arise. On the one hand, wants have to be dovetailed and awarded priorities. On the other, factors have to be combined in the best proportions. Usually plans are co-ordinated through numerous committees, directed at the top by a central planning committee whose members are primarily politicians.

Fourth, it is argued that state ownership of resources, by reducing personal incentives, diminishes effort and initiative. Direction of labour may mean that people are dissatisfied with their jobs; officials may play for safety in their policies (see p. 166). Thus production may be less than under private enterprise.

Fifth, and probably most important, there is the political danger. Once individuals have given power to the state to decide what is good for them and to direct

all factors of production, it may eventually seize absolute political power. Thus the ultimate choice between a market economy and a command economy (in their extreme forms) really hinges on whether people prefer to run the risk of dictatorship or to accept the defects of the market economy which allows them to choose their own jobs.

1.5 Britain's mixed economy

(a) The 'middle way'
Fortunately a community does not have to make a complete choice between these two extremes. Instead it can compromise by recognising that, while the private enterprise economy responds to individual preferences, the state has a role to play in overcoming its defects.

Thus, in an attempt to get the best of both worlds, the UK can be said to have a 'mixed economy' in which three-quarters of production is carried out by the private sector through the market (though subject to varying degrees of government control), while for the other quarter the government is directly responsible through the public sector. Moreover, chiefly by income redistribution and subsidies, the government influences the allocation of the goods and services produced.

(b) Summary of the objectives of government economic policy
The objectives of government economic policy in the UK fall under three broad headings: allocation of resources, stability of the economy and the distribution of income.

(i) *The allocation of resources*
Our earlier summary of the defects of the market economy, suggests that the government must be concerned with:

(a) imperfect competition, particularly with regard to monopoly;
(b) the immobility of resources;
(c) imperfect knowledge of both consumers and producers;
(d) community goods, where a pricing system cannot be operated;
(e) external costs and benefits, especially in the context of the environment.

(ii) *Stability*
Stability of the economy involves the government taking measures to achieve:

(a) full employment;
(b) a stable price level;
(c) a balanced regional development;
(d) a healthy balance of payments;
(e) a steady and acceptable rate of growth.

(iii) *The distribution of income*
Since satisfaction is personal to the individual (see p. 10), welfare resulting from a redistribution of income cannot be dealt with scientifically by objective

measurement (see below). However, many people do feel that poverty in the midst of plenty is unacceptable, and their satisfaction is increased by giving to voluntary agencies, e.g. the Salvation Army. Yet while voluntary bodies play a part in the redistribution of income, their efforts are limited and concentrated mostly on sectional interests. Only the government can achieve the degree and fairness of redistribution which is acceptable.

The government may carry out redistribution by deliberate policy measures (as with public spending and taxation). But redistribution may take place only as an offshoot of measures designed to achieve other objectives, e.g. equal opportunities legislation.

(c) Private and public sectors
In examining how a mixed economy works, it is convenient to distinguish between the 'private sector' and the 'public sector'. The former includes those firms which are privately owned and where decisions are taken in response to market signals. The latter includes government departments, local authorities, the nationalised industries, and public bodies such as the Universities Funding Council. All are distinguished by the fact that their capital is publicly owned and their policies can be influenced through the ultimate supply of funds by the government. Thus the existence of the public sector enables the government to exercise an important measure of control over the economy. Moreover, decisions on what to produce can be based on need rather than demand (see p. 164).

1.6   The nature of economics

(a) Welfare and wealth
The economist who is concerned with human welfare must recognise that in the last resort people do not want goods as such, but simply the satisfaction they obtain when consuming those goods. Yet it is quite impossible to measure satisfaction. It is probable, for instance, that a schoolgirl derives more enjoyment from £55 spent on a new tennis racket than a millionaire does from £55 spent on a dinner. Yet we can never be sure – satisfaction, like love or pain, is a personal feeling which cannot be measured objectively.

So the economist, working on the principle that two loaves are better than one, measures the output of goods and services and declares that any increase over a given period of time indicates an increase in welfare. Nevertheless, as we see in chapter 16, this is only an approximation, and its limitations have to be constantly borne in mind.

(b) The role of economics
It will be appreciated that in practice many economic decisions involve subjective judgements; that is, they cannot be made solely by an objective appraisal of the facts but depend to some extent on personal views in interpreting facts. Thus the relative size of the public sector and the extent to which the government interferes with the operations of firms in the private sector are determined largely by the political philosophy of the elected government.

The economist tries to be as objective as possible, establishing principles which, given certain conditions, show how the economy works and can be used to predict the likely results of policies.

Decision-makers may brush these principles to one side, either because facts necessary for a complete answer are not available or because different weight is given to assumptions. But at least economics provides a reminder of where objectivity ends and subjectivity begins.

② The population as consumers and producers

2.1 Growth of population

(a) Changes in population size
The people of a country are its consumers. They also provide the labour force for production. A study of the population of the UK, therefore, gives a bird's-eye view of the community for which the economic system must provide, and also of the size and nature of the available labour force.

At any one time the structure of the population is largely the result of demographic factors prevailing some fifty years earlier. It is necessary, therefore, to consider such factors in order to explain the UK's present pattern of population and how it is likely to change in the future. Owing to the difficulty of obtaining consistent figures for Northern Ireland, the discussion concentrates on Great Britain. The basic conclusions, however, apply equally to the UK as a whole.

Table 2.1 and Figure 2.1 reveal that, while the UK's population increased quite rapidly during the nineteenth century, there was a marked falling-off in the rate of increase in the twentieth century. This poses three main questions. (i) Why was there such a rapid growth during the nineteenth century? (ii) Why did that rate fall so markedly during the twentieth century? (iii) What is likely to happen during the rest of the twentieth century?

Table 2.1 Population (in 000s) Great Britain and Northern Ireland, 1801–1991

Date	Great Britain (England, Wales and Scotland)	Northern Ireland
1801	10 500	1 442
1851	20 817	1 237
1901	37 000	1 371
1951	48 854	1 536
1971	53 979	1 539
1981	54 286	1 490
1991	54 889	1 578

Source: *Annual Abstracts of Statistics*.
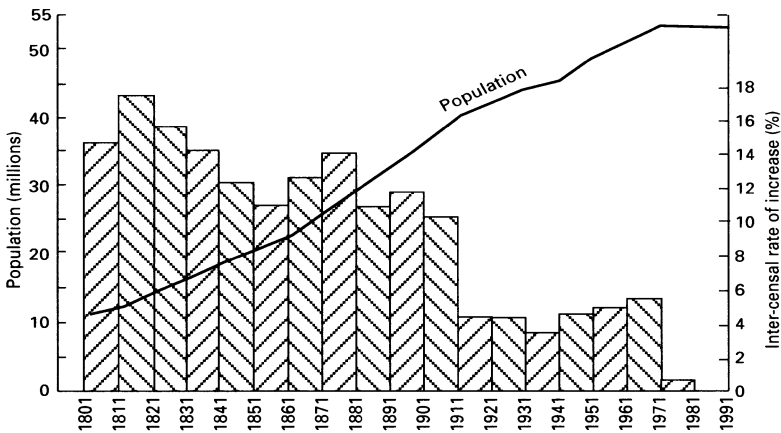
Figure 2.1   Population 1801–1991, Great Britain

(b) Causes of changes in the rate of growth

The factors affecting population changes are shown in Figure 2.2. On the one hand we have the natural increase – the excess of births over deaths; on the other, migration – the balance between immigration (inwards) and emigration (outwards). In fact, apart from the years 1931–41 and 1951–61, Great Britain has lost by migration about half a million people each decade. Changes in the rate of growth, therefore, have resulted chiefly from changes in the natural increase.

For our purposes crude rates are adequate for examining changes in the rate of births and deaths. The *crude birth rate* (CBR) is the number of births per year per thousand of the population. For example, if the total population is 50 million and the number of births in the year is 1 million, the CBR is 20. Similarly the *crude death rate* (CDR) is the number of deaths per year per thousand of the population.

The reason for the high rate of increase during the nineteenth century was that, while the birth rate remained high, the death rate fell considerably (probably from about 33 in the mid-eighteenth century). This fall was the result of
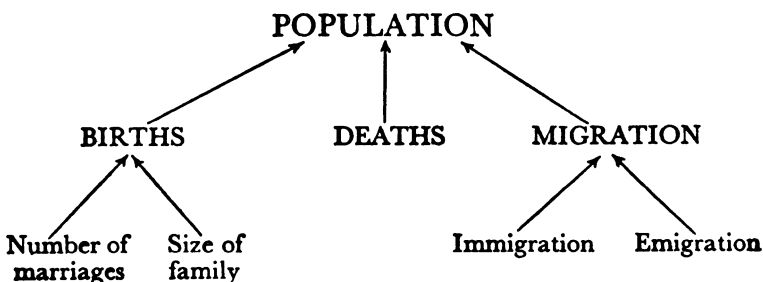
Figure 2.2   Factors affecting population

improved medical knowledge, better sanitation and water supplies, and the higher standard of living following the agricultural and industrial revolutions.

But the situation changed in the twentieth century. The death rate did not fall so rapidly. More important, the birth rate fell considerably. The reason was a decrease in the average size of family – from between five and six children to just over two. A variety of factors contributed to this: improved methods and social acceptance of birth control; the increased economic burden of parenthood – due, for instance, to the gradual raising of the school-leaving age; the higher standards which parents generally set themselves for their children's welfare; the growth of competing alternatives to children, such as holidays, foreign travel, the cinema and the motor-car; the emancipation of women, politically, economically and socially, with the consequent desire to be free from home ties; and the momentum which social example, smaller houses and advertisements provided once the movement towards smaller families had started. In 1949 there was a real possibility that Britain's population would have declined in number by the end of the century!

Today, however, such a decline seems less likely. Until recently, people were building slightly larger families, the result of younger marriages, greater economic prosperity, increased government help to families and improved opportunities for younger mothers to resume work.

But the 1991 census revealed that the rate of increase of the population over the previous ten years had dropped to only 1.2 per cent. While external emigration had been a contributory factor, this low rate of increase was mainly due to a fall in the size of family. Not only has the marriage rate fallen and people are getting married older, but they are leaving it longer after marriage before starting their families. On the other hand, the first child is now followed more quickly by the second, but those couples having 3 or more children has halved compared with 1971. This change in the pattern of family-building has been brought about largely by the desire of married women to shorten the period when they are not available for work.

Table 2.2   Crude birth rate (CBR) and crude death rate (CDR) 1851–1991, England and Wales

Date	CBR	CDR
1851	35.5	22.7
1900–2	28.7	17.3
1950–2	16.0	11.9
1961	17.9	12.0
1971	16.2	11.6
1981	13.0	11.8
1991	13.8	10.9

Source: *Annual Abstract of Statistics.*

As a result, the average size of family has fallen to around the 2.20 level, where the population would just be replacing itself.

While any projections of population depend upon the reliability of assumptions, especially as regards births and migration, it now seems likely that the population of the UK will be about 58 million at the end of the century. Whether such an increase is desirable or not will now be examined.

2.2   Implications of changes in size of population

(a) The Malthusian theory of overpopulation

Until the middle of the eighteenth century, the population of Britain grew slowly. But, from then on, growth became more rapid, and in 1798 Thomas Malthus's first essay on *The Principle of Population as it affects the future improvement of Society* made it a major subject of discussion.

Malthus began from two postulates: (i) that the passion between the sexes is necessary and will remain nearly in 'its present state'; and (ii) that food is necessary to the existence of man. Given these postulates, his arguments forced him to conclude that: (i) the population will, if unchecked, double itself every twenty-five years; and (ii) the means of subsistence can, at a maximum, increase by only the same amount every twenty-five years. In other words, while population multiplies in a geometric progression, food supplies increase in an arithmetic progression.

The first conclusion was based on information collected by Malthus on the populations of various countries. But the second was supported by no evidence whatsoever. In order to substantiate it, Malthus appealed to the 'known properties of land'. Here he was virtually relying on the *law of diminishing returns* (see p. 85), although that was not precisely stated until some fifty years later.

From these two conclusions the important result followed that the power of population to increase was 'infinitely greater than the power of the earth to produce subsistence for man'. In short, the population would always tend to outgrow its food supply.

Since man cannot live without food, what, Malthus asked, kept population within its means of subsistence? The answer he found in certain 'checks'. First, there were 'positive checks', involving misery – famine, war, disease. Second, there were 'preventive checks' – which, with one exception, all involved 'vice', including contraception. The exception was 'moral restraint', by which was meant deliberately refraining from marrying at an early age. Since this was a remote possibility, the outlook for civilisation was gloomy: in the long run mankind could only expect a subsistence level of existence. Moreover, social policies to alleviate poverty would be self-defeating.

(b) Malthus's 'blind spots'

Although at the beginning of the nineteenth century Malthus's views were widely accepted, the final tragedy of starvation, the logical outcome of his two conclusions, has not occurred. Where, therefore, did Malthus go wrong?

First, we must note that to some extent his argument was illogical, for he did not deal with the fact, well known at the time, that, in spite of the rapid increase in the population over the previous fifty years, people on the average were no worse off. This showed that the means of subsistence must at least have increased in proportion. Had Malthus possessed a precisely formulated law of diminishing returns, he could have based his argument on a fixed total supply of land which would sooner or later make itself felt as the population increased.

Second, Malthus was preoccupied with people as consumers. He failed to see that, by and large, a consumer is also a producer, for 'with every mouth God sends a pair of hands'. Here again a fixed supply of land with consequent diminishing returns could have overcome this objection.

Third, Malthus did not foresee change. On the one hand the geometric increase in Britain's population did not come about, because of emigration and above all because of the reduction in the size of the family with rising living standards. On the other hand, improved agricultural techniques and the vast increase in imports meant that Britain's food supplies were not limited to increasing in an arithmetic progression.

Thus Malthus's arguments have validity only when there are fixed resources, such as land or energy reserves. It is, for instance, the limited supply of land which brings about a Malthusian situation in the Far East today.

(c) The concept of an 'optimum population'
To Malthus, increasing numbers were a bad thing, as they pressed on the means of subsistence and lowered the standard of living. But his views lost ground towards the middle of the nineteenth century, for, as the capital investment of the industrial revolution began to yield benefits, the standard of living was seen to be keeping pace with the increase in population.

Indeed, at the turn of the century, Professor Edwin Cannan showed that population could be too small to take full advantage of available technical knowledge. For example, a large population might justify large-scale production, with more use being made of division of labour, specialised machines and technical discoveries. In short, a doubling of the population could lead to more than doubling production.

Since, therefore, population can either be too large or too small, there must be an intermediate point where it is just right: the optimum population at which, given existing technical knowledge, capital equipment and possibilities of exchange with other countries, average output per head is at a maximum. Thus, with reference to Table 6.1 on p. 85, the optimum population for the example given would be four labourers. It follows that any country is over- or underpopulated if its population is respectively more or less than the optimum.

But the concept of an optimum population is not without difficulties. In the first place it is unjustifiable to specify 'given existing technical knowledge, capital equipment and possibilities of exchange' and then to speculate as to what production would be if the population were larger or smaller. Had the population increased differently, these variables themselves would have been different. The same mistake is apparent in J.S. Mill's argument in the middle of the nineteenth

century that the world would have been better off if, with the improvements that had taken place, population growth had been less. The truth is that such improvements would not have taken place, for a large and rapidly growing population accumulates knowledge and equipment differently from a smaller or slowly growing one. Even more important is that, from the practical point of view, the concept is of little help. Any optimum population at which a country was aiming would only remain the optimum so long as technical knowledge, etc., did not change. Thus, before an optimum was achieved, some new figure would have taken its place. All that can be done, therefore, is to consider the present composition of the population, forecast the population which will result from it, and then relate this population to likely changes in capital accumulation and technical discoveries.

   We now apply this procedure to a study of Britain's population.

(d) The advantage to Great Britain of an increasing population
An increasing population has certain advantages which stimulate growth:

(i) *It increases the size of the home market*
The additional output needed for a larger population should benefit industries working under conditions of decreasing costs, e.g. those producing aircraft, computers, nuclear reactors. It should be noted, however, that this applies only if the extra output is provided by existing firms and not by additional firms entering the industry. Moreover, it is possible to obtain large-scale economies by specialising on a narrow range of goods and exporting, as in Switzerland.

(ii) *It facilitates labour mobility*
With an increasing population, unemployment resulting from the immobility of labour presents fewer problems. This is because the decline of older industries is slower and can be covered by natural wastage with fewer redundancies, while expanding industries can obtain most of their additional workers from new entrants to the labour force.

(iii) *It encourages investment*
An increasing population makes it easier to maintain the level of replacement investment. More than that, the extra consumer demand necessitates additional investment in machinery, factories, schools, houses, transport, etc. Consequently it stimulates improved techniques, thereby accelerating the replacement of existing equipment.

(iv) *It promotes vitality*
By weighting the age distribution in favour of youth, an increasing population ensures more workers per retired persons and makes for energy, mobility, inventiveness and willingness to accept new ideas.

   It should be noted that the disadvantages of a decreasing population could be stated as the opposite of these.

(e) The disadvantages of an increasing population

Against the advantages given above it is necessary to set certain disadvantages which may make it difficult for an increasing population to raise present living standards. Resources have to be used in adding to capital equipment instead of in producing consumer goods or improving existing buildings. Moreover, the saying that 'with every mouth God sends a pair of hands' ignores two important facts. The first is that not every person is a producer – for a time the additional mouths have to be provided with food, education, etc., by the working group. The second is more important: the increase in the number of labourers on a fixed amount of land may well bring the law of diminishing returns into operation, with a consequent fall in living standards.

It is the law of diminishing returns which pinpoints the problem of increasing numbers. In Africa and Asia there is simply a lower output per head, as extra people have to obtain their subsistence from a fixed amount of land. But for Britain the law of diminishing returns does not apply in this basic way. The law assumes no technical improvements, but improved techniques in agriculture have enabled Britain in recent years to raise her self-sufficiency in her principal foodstuffs from 50 per cent to 90 per cent.

As far as Britain is concerned the law of diminishing returns is significant in three ways. First, Britain is still very dependent on imports for the raw materials, special types of machinery, base metals, consumer goods (e.g. cars, cameras) and certain sorts of food and drink (e.g. fruit, wine) that she requires. What Britain has to ask, therefore, is: can exports be increased sufficiently to pay for the extra imports required by the larger population without a severe deterioration in her terms of trade especially when North Sea oil eventually runs out?

Throughout the nineteenth century Britain proved that this was possible. Indeed a balance-of-payments surplus allowed her to invest heavily abroad. But since then the problem of finding and holding foreign markets has become more difficult.

Second, the growth of population accompanied by increasing real income puts pressure on space for amenity and recreational purposes, as, for example, more houses with spacious gardens are demanded and access to the countryside is sought.

Third, with population growth, environmental problems intensify. As city congestion increases and more open space is required for housing, roads and industry, arguments for conservation and control of pollution gain momentum.

(f) Conclusion

At present Britain is managing to support increasing numbers while improving living standards. But the overall level of the future population must be watched by the government and, if necessary, influenced by immigration policy and the level of child benefits. Social as well as economic considerations have to be taken into account.

In Africa and Asia a Malthusian situation exists. While the death rate is falling through better medical services, the birth rate remains high. By the end of the century an almost doubled population could be seeking to live with little increase

in land. Possible solutions are birth control, improved agricultural techniques, the development of export industries, and economic aid from developed countries.

2.3   Age distribution

Any change in the birth or death rates will affect the age distribution over time. The present overall pattern is of an ageing population brought about by the fall in the rate of increase during the twentieth century (Table 2.3).

(a) The effects of an ageing population
It should be noted that in part this trend is due to a normal development, the fall in mortality. This means that we have to adjust to the change: trying to prevent it by raising the birth rate will also increase the size of the population – which, as we have seen, presents problems.

It is essential, therefore, to anticipate the possible economic, social and political difficulties resulting from changes in the age structure so that the necessary adjustments can, with planning, occur smoothly.

(i) *Economic*
First, there is an increased dependence of retired persons on the working population. Current wants can only be provided for by current production. An ageing population means that the proportion of workers to consumers is falling. Whereas in 1851 there were over 12 people of working age for every person over 65, in 1991 there were only 4. Indeed the number of pensioners is likely to increase from 10.5 mn in 1991 to 14.6 mn in 2031 with only 3 workers to support each pensioner. This highlights the importance of economic growth.

Second, a changing pattern of consumption will result. An ageing population means, to take extreme examples, that wheel chairs will be wanted in place of prams, walking-sticks in place of hockey-sticks, tea in place of milk. For many of these new wants, consideration has to be given well in advance. Today almost

Table 2.3   Changes in the age distribution of the population of the UK 1961–91 and 2001, 2021 (projected)

| | *(% of population)* | | |
	under 16	*16–64*	*65+*
1961	24.8	63.4	11.8
1971	25.6	61.2	13.2
1981	22.2	62.8	15.0
1991	20.3	63.9	15.7
2001 (*projected*)	21.3	63.1	15.6
2021 (*projected*)	19.5	62.4	18.1

Source: *Social Trends*, 1993.

one-quarter of British households consist solely of pensioners and, when plan-
ning a housing programme, we have to provide the smaller units and sheltered
housing they require. Similarly, the fall in births since 1964 has resulted in fewer
school places and teachers being required.

Third, older people make greater demands on the health service, those over 75
years, for instance, costing four times as much per head as the average person.
Thus national health expenditure has to be increased and diverted towards the
medical care required by older people, e.g. hip operations. Eventually more of
this growing need may have to be met through private health insurance.

Fourth, an older labour force is less mobile. In the past, expanding industries
have obtained labour from young people just starting their working lives, while
the decaying industries have declined fairly quietly by natural wastage – not
replacing workers as they leave or retire. However, where the working popula-
tion is static in size, expanding industries have to draw older workers from the
declining industries. 'Teaching old dogs new tricks' and moving them to new
areas is not always easy (see pp. 136–8). A high level of unemployment increases
the difficulties, for it is the older workers in the declining industries who are
likely to remain out of a job the longest. Thus both the government and firms
must provide training schemes and re-location incentives.

Fifth, an older population tends to be less progressive. While older people
are more patient and experienced than younger people, the latter excel in
energy, enterprise, enthusiasm and the ability to adapt themselves and to learn
new skills.

(ii) *Social*

Where old people are more numerous and live longer, their children find greater
difficulty in caring for them. Thus there is an increasing need for the state to pro-
vide home-care services (such as 'meals-on-wheels' and home helps) and old
people's homes.

Similarly, there is a greater demand for advice from Citizen's Advice Bureaux,
since older people need more help in sorting out difficulties relating to housing,
gas and electricity bills, social-security benefits and so on.

(iii) *Political*

Political decisions have to be made as a result of the disadvantages of an ageing
population. To what extent should younger generations be augmented by a liberal
immigration policy, bearing in mind the social stresses which could arise? Can
adequate defence be provided by the use of more sophisticated weapons, or will
the falling proportion of young people necessitate conscription? Should TV and
radio programmes give greater weight to the type of entertainment preferred by
old people? And, since older people own the larger share of the nation's capital
but are adverse to taking risks, should the state assume responsibility for provid-
ing funds for the riskier types of enterprise, such as North-Sea-oil prospecting
and development and the exploration of outer space?

Indeed, the ageing population may even influence the election of governments,
for older people tend to be more conservative.

2.4   The industrial distribution of the working population

(a) The working population

The UK's population in 1991 was 56,467,000 persons. Of these, 28,340,000 persons (16,240,000 males and 12,100,000 females) are described by the Department of Employment as 'the working population'.

The working population consists of persons over school-leaving age who work for pay or gain or are claiming unemployment benefit. It therefore includes both employees and self-employed persons, even if they are over retirement age or are working only part-time, and members of the armed forces. Excluded are: (*a*) children under 16 years of age and students above 16 years of age who are receiving full-time education; (*b*) persons, such as housewives, who do not work for pay or gain; (*c*) persons who, having private means, e.g. from investments or gifts, do not need to work; (*d*) retired persons; (*e*) work-seekers who do not claim benefit, e.g. housewives who are ineligible.

The size of the working population depends upon:

   (i) The numbers within the 16–65 age-group.
  (ii) The activity rates within this group, especially as regards young people and female workers. The tendency over the last twenty years has been for a higher proportion of young people to remain in further education, thus reducing their activity rate. On the other hand, a higher proportion of women are now entering the working population. The expansion of the service and light-manufacturing industries has provided increased job opportunities for women, while the changed attitude to women workers is reflected in the Equal Pay Act 1970 and the Sex Discrimination Act 1975. Above all, the smaller family, the development of part-time employment opportunities and new labour-saving domestic appliances have allowed married women to work.
 (iii) The extent to which people over retiring age continue to work, something which is largely influenced by the level of pensions.

Table 2.4   Total working population, UK, 1975–91 (000s)

	1975	*1980*	*1991*
Total employees in employment	22 710	22 458	22 259
HM Forces	336	323	297
Self-employed persons	1 993	1 950	3 143
Work-related government training programmes	–	–	343
Unemployed	838	1 444	2 293
Total working population	25 877	26 839	28 340
of whom: Males	16 162	16 320	16 240
Females	9 715	10 520	12 100

Source: *Annual Abstract of Statistics.*

The main changes in the working population since 1975 are:

(1) An overall increase of nearly $2^1/_2$ million (9.5 per cent). Thus to day the working population covers one-half of the total population.

(2) Of this increase, 2 million have been in female workers. Here the main cause has been the expansion in the service industries. Forty per cent of female workers are part-time (compared with only 5 per cent males). In the UK, the activity rate of women of 16–60 years of age is 66 per cent, higher than all other countries except for Japan and the Scandinavian countries.

(3) Self-employment has increased by 1,150,000 (58 per cent), largely through the active support of the government. It occurs mostly in agriculture, construction, distribution, catering and other services.

(4) As regards occupation, the movement has been from traditional craft and semi-skilled jobs towards more highly skilled technicians and technologists.

(b) Changes in the industrial distribution of the working population

Table 2.5 shows significant changes in the industrial distribution of employees between 1985 and 1991:

(1) a 15 per cent decrease in agriculture, forestry and fishing;

(2) a 15 per cent increase in services;

(3) a 10 per cent decrease in manufacturing industries.

The basic explanation of the first two changes is an increase in real income over the period, and in this respect it is merely a reflection of a continuous trend. As a result, spending moves to those goods having a high income elasticity of demand. In 1985, for example, the average worker spent 25 per cent of his income on food; by 1991 it had fallen to 15 per cent, of which one-third went on

Table 2.5   Employees in employment by industry: UK, 1985–91

	1985		*1991*	
	Total (000s)	*Per cent*	*Total (000s)*	*Per cent*
Agriculture, forestry and fishing	341	1.6	291	1.3
Energy and water supply	591	2.8	439	2.0
Manufacturing industries	5 362	25.0	4 822	21.7
Construction	1 021	4.8	962	4.3
Distribution, hotels, catering and repairs	4 295	20.0	4 686	21.0
Transport and communication	1 327	6.2	1 349	6.0
Banking, insurance and other financial services	2 068	9.6	2 693	12.1
Other services	6 418	30.0	7 026	31.6
All industries and services	21 423	100.0	22 268	100.0

Source: *Annual Abstract of Statistics.*

meals bought outside the home. Today the provision of services now covers three-fifths of all employees.

The decrease in manufacturing has largely been the result of the uncompetitiveness of British products in world markets brought about by: (*a*) the greater technical efficiency of Japan; (*b*) the inability to contain wage costs relative to competitors; (*c*) the appreciation in the value of the £ sterling after 1979 as the result of the export of North Sea oil.

Technical advances and increased mechanisation also help to explain the relative falls in agriculture, energy supply, construction, transport and communications (e.g. increased car ownership), and manufacturing. In contrast, services are more labour-intensive and less highly-capitalised.

2.5   The geographical distribution of the population

Geographically, the population of the UK is dominated by two features: it is *concentrated*, and it is *urban*. Both are the result of moving from an agricultural to an industrial economy.

(a) The concentrated nature of the population
As a result of the industrial revolution industry migrated to the coalfields in the Midlands and north of England. And today, even though electricity frees industry from being located on the coalfields and the basic industries of these areas have declined, they still remain important centres of industry and population (Figure 2.3). There are two main reasons for this. First, many industries remain because of acquired advantages, particularly the availability of labour (see p. 76). Second, new industries have been attracted by the government's Development Area policy (see p. 256).

During the twentieth century, the main areas of natural expansion have been the Midlands and south-east England, particularly in the counties around London. The Midlands expanded with the demand for light engineering and electrical products and for motor-vehicles, but the area has recently suffered some stagnation with the decline in the motor and other manufacturing industries. On the other hand, London, always a relatively wealthy area, has expanded outwards as production of goods having a high income-elasticity of demand has moved nearer to their main markets. In addition, the development of government and financial services has attracted population, and this tends to have a multiplier effect as other industries move in to provide for their needs. The proximity of south-east England to the Continent has also attracted industry and population, especially since 1973 when Britain joined the EC.

Certain rural counties, although not of high density, have also shown a high rate of growth, particularly Norfolk, Suffolk and Cornwall.

(b) The urban nature of the population
This concentration of population is in towns (unlike the Nile and Ganges deltas which consist mainly of concentrated rural communities). Over three out of every four persons in Britain live in an urban area. Moreover, 30 per cent of the

Figure 2.3   Geographical distribution of the population of the UK

population lives in the seven conurbations (continuous built-up areas) of Greater London, south-east Lancashire (Manchester), west Midlands (Birmingham), central Clydeside (Glasgow), west Yorkshire (Leeds and Bradford), Merseyside (Liverpool) and Tyneside (Newcastle).

Nevertheless, over the last thirty years, the proportion of the total population in rural areas has increased, reversing the trend towards town life of the previous 150 years. As cities grow, retailing and commerce compete for central sites. The

consequent rise in land values pushes the residential population outwards. More-over, as real income increases, people can afford houses with more space in the suburbs. Thus *inner*-city areas have been losing population.

Indeed, in the conurbations the population is moving beyond the green belt into rural areas. Long-term factors are bringing this about: (*a*) the growth in real income which covers the higher costs of transport; (*b*) the mobility provided by the car and faster rail travel; (*c*) the improvement in public utility services; (*d*) the direct stimulus of the new towns policy; (*e*) the expansion of new light industries in rural towns; (*f*) a greater appreciation of rural life.

(c) Problems resulting from the concentration of population in urban areas
The concentration of population in large urban areas has certain advantages. Such areas can offer better and more specialised schools, shops, entertainment and other services. Fast road, rail and air communications facilitate travel between cities. They can usually provide a variety of employment opportunities, enabling firms to recruit the different types of labour they require. But conurba-tions often involve travelling long distances to work within the urban area, put-ting a heavy stress on the transport system. There are also problems of inner-city decay, with poor housing, inadequate schools, pollution and lack of open spaces. Moreover, there are fewer social ties, and the lower community spirit results in vandalism and petty crime. Above all, where these areas are faced with the decline of major local industries, there is the problem of regional unemployment. Not only do government organisations have to respond to these conditions, but the development of urban areas may itself lead to the reorganisation of local government.
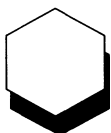
Part II
Private-sector production

③ The operation of the free market

3.1 Markets

(a) Value and price
In the market economy a want is significant only when a person is prepared to give up something in order to satisfy it. As the strength of the different wants varies, so will the amounts which people are willing to give up. In other words, different goods have different *values*, value being the *rate* at which a particular good or service will exchange for others. In modern economic systems the values of goods are expressed in terms of money, as *prices*.

Changes in *relative* prices, if supply conditions have not changed, indicate a relative shift in the importance of the goods concerned. Thus price changes signal changes in what people want. We must therefore examine the mechanism by which these signals are flashed up. We begin by looking at the *market*.

(b) What is a market?
'I am offered £1050 for this heifer. No more offers? For the last time of asking, any advance on £1050? Going at £1050, going, gone.' Down comes the hammer. 'Sold at £1050 to Mr Giles on my right.'

This is the local cattle-market. On his stand above the cattle-ring is the auctioneer. Inside the ring, a black-and-white heifer is appraised by local farmers and dealers. Some are buyers, some sellers. The market fixes the price at which those who want something can obtain it from those who have it to sell.

Note that it is only exchange value which is significant here. The farmer selling the heifer may have felt that it ought to have made more than £1050. Or, as it was the first calf reared by his son, it may have had great 'sentimental value' to him. Such considerations, however, mean little in the market economy.

Of course, prices are not always fixed by auction. This is the method usually employed where there are many buyers but the seller only comes to the market infrequently, or wishes to dispose of his goods quickly. If there are few buyers and sellers, e.g. in the purchase of a house or a second-hand car, the final price may be arrived at by 'higgling' – the seller meeting the prospective buyer personally and bargaining with him.

But where goods are in constant demand the above methods take too long. Thus most goods, such as foodstuffs, clothing and household utensils are given a definite price by the shopkeeper. But buyers will still influence this price. If it is too high, the market will not be cleared; if it is too low, the shopkeeper's stocks will run out.

A market need not be formal or held in a particular place. Secondhand cars are often bought and sold through newspaper advertisements. Secondhand furniture may be disposed of by a card in a local shop window. Foreign currency, gold, base metals, raw cotton and other goods which can be accurately described are dealt in over the telephone.

However, in studying the market economy it is essential to understand how price is determined. Since this is done in the market, we can define the market simply as *all those buyers and sellers of a good who influence its price*. Within the market there is a tendency for the same price, allowing for costs of transport, to be established for the same commodity.

(c) World markets

Today modern transport allows many commodities to have a world market – a price change in one part of the world affects the price in the rest of the world. Examples of such commodities are wheat, coffee, oils, basic raw materials (such as cotton and rubber), gold, silver and base metals. What conditions must a commodity fulfil to obtain a world market?

First, there must be a wide demand. The basic necessities of life (e.g. wheat, vegetable oils, wool, cotton) answer this requirement. In contrast, such goods as national costumes, books translated into little-used languages and postcards of local views have only a local demand.

Second, commodities must be capable of being transported. Land and buildings are almost impossible to transport. Personal services are limited by the distance the consumer can travel. Labour, too, is particularly immobile, especially when it comes to moving to a different country (see Chapter 10). Furthermore, governments may, by import taxes and quotas, effectively prevent the entry of certain commodities into the country.

Third, the costs of transport must be small in relation to the value of the commodity. Thus the market for diamonds is worldwide, whereas that for bricks is local. Similarly wheat and oil are cheap to transport compared with coal because they are more easily handled – although, as sea transport is relatively cheap, coal mined near the coast can be sent long distances.

Last, the commodity must be durable. Goods which perish quickly, such as milk, bread, fresh cream and strawberries, cannot be sent long distances. Nevertheless modern developments, such as refrigeration, canning and air freight transport, are extending the market even for these goods.

(d) Perfect and imperfect markets

In any market the price of the commodity in one part affects its price in another part. Hence the same price tends to be established. Where price differences are eliminated quickly, we say the market is a 'perfect' market. (Note that this is not quite the same as 'perfect competition' – see chapter 6).

For a market to be perfect certain conditions have to be fulfilled. First, buyers and sellers must have exact knowledge of the prices being paid elsewhere in the market. The development of communications, particularly telecommunications, has facilitated this. Second, both buyers and sellers must base their actions solely

on price, and not favour one particular person out of loyalty or mere inertia. Thus, if one seller puts up the price of his good, his customers immediately go to another who is cheaper. Alternatively, if he lowers his price, customers will so flock to him that he would sell out quickly unless he raises his prices to that asked elsewhere.

Examples of perfect markets are the precious-stones market of Hatton Garden in London, and above all the organised produce markets and the stock exchange (see below). In these markets the two essential conditions are fulfilled, for prices are watched closely by professional dealers. As a result of their operations, variations in price are quickly eliminated.

But such conditions are rarely satisfied in other markets. Buyers and sellers neither have perfect knowledge nor act solely on the basis of price. The ordinary housewife, for instance, cannot afford the time to go from one shop to another in order to compare the prices of her everyday purchases, though she is usually much more careful when spending on the more expensive goods bought at infrequent intervals. Similarly, shopkeepers do not always know what competing shopkeepers are charging for their goods. Moreover, purchasers may be influenced by considerations other than price. Thus they may continue to deal with one particular trader, even though he is charging a slightly higher price, because he has given them good service in the past. Finally, although two goods may be virtually the same physically, by 'product differentiation' and advertising the merits of his own brand a producer may convince the consumer of its superiority. Such 'persuasive' advertising, which accounts for over a half of present advertising expenditure, makes the market less perfect, and must be contrasted with 'informative' advertising, which increases knowledge and thus helps to make the market more perfect.

Where price differences persist, markets are said to be 'imperfect'. As we have already hinted, such markets are often found in retailing.

(e) Organised produce markets

As explained above, the market for certain commodities is worldwide. Moreover, many of these commodities are in constant demand, either as basic raw materials or as main foods or beverages for a large section of the world's people. They therefore figure prominently in international trade, and are the subject of the following discussion.

England's foreign trade began with the export of raw wool in the thirteenth century, and it was extended by the subsequent development of the chartered companies. These were based in London, and it was there that merchants gathered to buy and sell the produce which the companies' ships brought from abroad.

The big change, however, came about with the expansion of international trade following the industrial revolution. The UK became the greatest importing and exporting nation in the world. London, her chief port and commercial city, not only imported the goods which were required for the people of her own country but, assisted by the fact that British ships were the carriers of world trade, built up an important entrepôt business, acting as a go-between in the distribution of

such commodities as tea, sugar, hides, skins and wool to many other countries, particularly those of western Europe.

Hence formal 'organised markets' developed. These markets are distinctive in that buying and selling takes place in a recognised building, business is governed by agreed rules and conventions, and often only certain persons are allowed to engage in transactions. They are thus a highly developed form of market. Today London has exchanges or auction centres for such commodities as rubber, wool, tea, coffee, furs, metals (tin, copper, lead and zinc), grain, and shipping freights (the Baltic Exchange). It must not be thought, however, that such organised produce markets exist only in London. Liverpool has exchanges for cotton and grain, and most of the other large trading countries have exchanges too. Although today many of the goods go directly to other countries, the earnings of London dealers are part of the UK's income from 'invisible exports' (see p. 295).

Broadly speaking, organised markets fulfil three main functions. First, they enable manufacturers and wholesalers to obtain supplies of commodities easily, quickly and at the competitive market price. Because they are composed of specialist buyers and sellers, prices are sensitive to any change in demand and supply. Thus they are perfect markets.

Second, 'futures' dealings on these markets enable people to protect themselves from heavy losses through price changes. Thus a cotton grower prefers to know what price he will receive before his output is actually delivered to the market. On the other hand a cotton spinner has to protect himself from a rise in the price of raw cotton between the time he quotes a price for his yarn and the time of its actual manufacture. Where a good is bought today for delivery today, the deal is known as a 'spot' transaction and the price is the 'spot price'. With many goods, however, it is possible to buy today for delivery in the future. The good may not even be in stock, but the seller contracts to obtain and deliver it at the agreed time. The price agreed upon is the 'future' or 'forward' price. For a commodity to be dealt in on a futures market certain conditions must be fulfilled: (i) the commodity must be durable, thereby enabling stocks to be carried; (ii) the commodity must be described in terms of grades which are internationally uniform; (iii) dealings must be frequent enough to occupy professional dealers; and (iv) the commodity must be subject to price fluctuations.

In future dealings the dealer uses his expert knowledge to make a profit on what he considers will be the future price of the commodity. At any time a dealer will quote a price (according to the view he takes of the future movement of prices) at which he is prepared to buy or sell at some future date. Thus a cotton grower can cover himself against a possible fall in price by selling his produce forward, while a cotton spinner can quote a weaver a price for yarn and guard himself against loss by buying the raw cotton forward.

Such dealing usually performs the third function of organised markets – evening out price fluctuations. At a time when an increase in supply would cause the price to fall considerably, the dealer adds his demand to the normal demand in order to build up his stocks, and thereby keeps the price up. On the other hand, when the good is in short supply he releases stocks and so prevents a violent rise in price. The difficulty is that speculation on the future price may dominate the

real forces which influence it, prices fluctuating violently in response to changes in optimism and pessimism.

3.2   Forces determining price

(a) Demand and supply

'That animal was cheap', remarks Phil Archer as the auctioneer's hammer falls. 'And no wonder', replies Brian Aldridge. 'This has been a long winter. We're now in the middle of April, and the grass is hardly growing. Hay's running short, and breeders are being forced to sell sooner than expected. Old Giles is about the only farmer who'll take the risk of buying extra cattle.'

What can we learn from Brian Aldridge's observations? Simply that the £1050 at which the heifer was sold was not really determined by the final bid. The real factors producing the relatively low price were the reluctance of farmers to buy and the number of young animals being offered for sale. In short, the price was determined by the interaction of the forces of demand and supply. We shall look at each in turn.

(b) Preliminary assumptions

First, we examine how these forces work in an imaginary market – for eggs. To simplify, we shall assume that:
  (i)   All eggs are exactly the same in size and quality.
  (ii) There are no transport costs within the market.
  (iii) The market consists of so many small buyers and sellers that there is keen competition.
  (iv) It is a perfect market – price differences are quickly eliminated because buyers and sellers (1) have complete knowledge of prices and conditions in other parts of the market, and (2) act solely on the basis of price.
  (v)   There is no interference by the government in the operation of market forces, e.g. by price control, regulating supply, etc.

(c) Demand

Demand in economics is the desire for something *plus* the willingness and ability to pay a certain price in order to possess it. More specifically, it is how much of a good people in the market will buy at a given price over a certain period of time.

It is helpful if we separate the factors affecting demand into (i) price, and (ii) the conditions of demand.

(i) *Price (the conditions of demand remaining unchanged)*

Normally a person will demand more of a good the lower its price. This is because, once you have some units of a good, you have partly satisfied your want and so will only buy more at a lower price. This conforms to our everyday observations. 'Winter sale, prices slashed' announce the shops when they wish to clear their stocks of clothing.

We can draw up a table showing how many eggs a person would be willing to buy at different prices. If they are very expensive, other foodstuffs will, as far as possible, be substituted; if they are cheap, people may even buy them to pickle. By adding up the demand from all buyers of eggs in the market at different prices over a given period of time, it is possible to obtain a *market demand schedule* (Table 3.1).

Table 3.1   Demand schedule for No Such Market for the week ending 29 January 1994

Price (*pence per egg*)	*Eggs demanded* (*thousands*)
12	3
10	9
8	15
6	20
4	25
2	35

*what buyers would take at each price.

Note that this schedule does not tell us anything about the actual market price or how much is in fact sold. It is an 'if' schedule. All it says is: '*If* the price is so much, then this quantity will be demanded.' Plotting this schedule on a graph, and assuming that demand can be obtained for intermediary prices, gives the demand curve *D* in Figure 3.1.
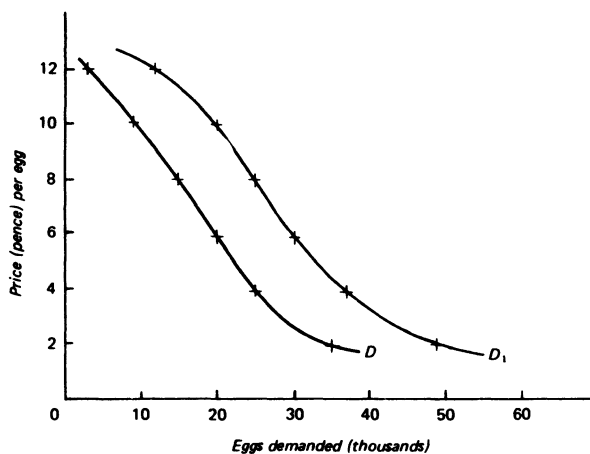
Figure 3.1   Quantity demanded and price

(ii) *The conditions of demand*
Something may occur to cause housewives to demand more eggs at a given price. In other words, the demand schedule alters. Suppose, for instance, farmers unite in an advertising campaign describing tasty egg dishes. As a result more eggs are demanded at all prices (Table 3.2).

Table 3.2   An increase in demand

Price (*pence per egg*)	Eggs demanded (*thousands*)
12	12
10	20
8	25
6	30
4	37
2	49

Plotting this revised demand schedule gives curve D_1 to the right of D. Had conditions so changed that demand decreased, the new demand curve would have been to the left.

The influence of both (i) price, and (ii) the conditions of demand, on the quantity demanded is thus shown on the graph. The former determines the shape of the demand curve – its slope downwards from left to right. The latter determines the position between the axes – an increase in demand shifting the curve to the right, a decrease to the left. For clarity's sake, a change in demand resulting from a change in price will in future be referred to as an *extension* or *contraction* of demand; a change in demand due to new conditions of demand will be described as an *increase* or *decrease* in demand.

Conditions of demand may change in a *short* period of time through:

(1) A CHANGE IN THE PRICE OF OTHER GOODS
Goods compete for our limited income and are thus, to some extent, substitutes for each other. When the prices of other goods fall, the particular good under discussion becomes relatively dearer and therefore less of it is demanded. When the prices of other goods rise, it becomes relatively cheaper, and so more of it is demanded.

But the effect on the demand for a particular good is more pronounced when the price of a close *substitute* changes. Suppose that fried tomatoes are an alternative to eggs for breakfast. If the price of tomatoes falls, housewives will tend to buy them rather than eggs. Thus, although there has been no initial increase in the price of eggs, demand for them has decreased. Similarly, where goods are *complements*, a change in the price of one good has a pronounced effect on the demand for the other. For example, a fall in the price of cars results in more cars

being purchased, leading eventually to an increase in the demand for petrol and tyres.

(2) A CHANGE IN TASTES AND FASHION
A campaign advertising eggs would increase demand; a scare that eggs were the source of infection would decrease it.

(3) EXPECTATIONS OF FUTURE PRICE CHANGES OR SHORTAGES
The fear that the price of eggs may rise considerably next week will induce people to increase their demand now in order to have eggs in stock.

(4) GOVERNMENT POLICY
A selective tax on eggs paid by the consumer would raise the price and lead to a decrease in demand; a rebate paid to the consumer would have the opposite effect (see p. 47 and pp. 273–6).
   Over a *longer* period the conditions of demand may change through:

(5) A CHANGE IN REAL INCOME
If there were an all-round increase in real income (that is, money income adjusted for any change in the price level) people could afford more eggs, and demand would probably increase. Or it might now be possible to afford mushrooms for breakfast, and these would take the place of eggs.

(6) GREATER EQUALITY IN THE DISTRIBUTION OF WEALTH
The wealth of a country may be so distributed that there are a few exceptionally rich people whereas the remainder are exceedingly poor. If many poor people felt they could not afford eggs, greater equality of wealth would be likely to increase the demand for eggs.

(7) A CHANGE IN THE SIZE AND COMPOSITION OF THE POPULATION
Additional people coming into the market will, with their extra income, increase demand, especially if eggs figure prominently in their diets.

(d) Supply
Supply in economics refers to how much of a good will be offered for sale at a given price over a given period of time. As with demand, this quantity depends on (i) the price of the good, and (ii) the conditions of supply.

(i) *Price (the conditions of supply remaining unchanged)*
Normally more of a good will be supplied the higher its price. The real reason for this is explained in Chapter 4. But even a brief consideration of how the individual farmer reacts to a change in price will show that it is likely to be true. If the price of eggs is high, he will probably consume fewer himself in order to send as many as possible to market. Moreover, the higher price allows him to give his chickens more food so that they can lay a few extra eggs. When we extend our analysis to the market supply it is obvious that a higher price for eggs enables farmers – including the less efficient – to produce more.

Hence we are able to draw up a *market supply schedule* – the total number of eggs supplied at different prices by all the sellers in the market over a given period of time (Table 3.3).

Table 3.3   Supply schedule for No Such Market for the week ending 29 January 1994

Price *(pence per egg)*	*Eggs demanded* *(thousands)*
12	40
10	32
8	25
6	20
4	13
2	7

*what sellers would offer at each price.

Once again it must be noted that this is an 'if' schedule, for all it says is: '*If* the price is so much, then this quantity will be offered for sale.'

We can plot this schedule (Figure 3.2); assuming supply for all intermediate prices, a supply curve S is obtained.
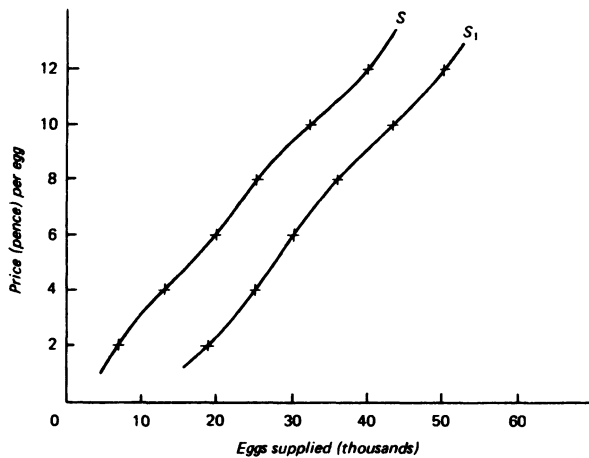
Figure 3.2   Quantity supplied and price

However there is a fundamental difference between demand and supply. Whereas demand can respond almost immediately to a change in price, a period

of time must usually elapse before supply can be fully adjusted. For the first day or two the only way in which the farmer can send more eggs to market is by eating fewer himself. By the end of the week he may have increased output by giving the hens more food or by leaving the light on in the hen-house all night; the higher price covers the extra cost. But to obtain any sizeable increase the farmer must add to his hens; if all farmers are following the same policy, this will take about five months, the period required to rear laying hens from chicks.

These different periods of time are dealt with more fully in chapter 6.

(ii) *The conditions of supply*

The number of eggs supplied may change even though there has been no alteration in the price. In the spring, for instance, chickens lay more eggs than in winter. Thus more eggs will be supplied at all prices in the spring, and fewer in winter, as shown in Table 3.4.

Table 3.4   An increase in supply in the Spring

Price (pence per egg)	Eggs supplied (thousands)
12	50
10	43
8	36
6	30
4	25
2	19

Table 3.4 shows that, compared with winter when only 25,000 eggs were supplied at 8p each (Table 3.2), during the spring 36,000 were supplied. Or, looked at in another way, 25,000 eggs can be supplied in the spring at 4p each compared with 8p in the winter. When plotted, the revised supply schedule gives a curve S_1 to the right of the old one. Had supply decreased, the new supply curve would have been to the left.

Like demand, therefore, supply is influenced by both (i) price, and (ii) the conditions of supply. The former determines the shape of the curve – its upward slope from left to right. The latter determines its position between the axes – an increase in supply shifts the curve to the right, a decrease to the left. To distinguish between the two we shall refer to a change in supply resulting from a change in the price of a commodity as an *extension* or *contraction* of supply; a change in supply due to new conditions of supply will be described as an *increase* or *decrease* in supply.

In general, conditions of supply may change fairly quickly through:

(1) PRICE EXPECTATIONS

Where a commodity is durable and the relative cost of storage low, e.g. gold, wheat, antiques, price expectations can, as with consumers' demand, affect supply. Thus if the price is expected to rise, stocks will be held or even augmented. If the price is expected to fall, stocks will be depleted. Supplies of perishable goods such as eggs are at a disadvantage here.

(2) A CHANGE IN THE PRICES OF OTHER GOODS, ESPECIALLY WHEN IT IS EASY TO SHIFT RESOURCES AND THE GOOD IN QUESTION IS IN 'JOINT SUPPLY'

Suppose, for instance, that there is a considerable increase in the price of chicken meat, including boiling fowls. It may pay the farmer to cull more of his older hens. Thus fewer eggs are supplied at the old price.

(3) A CHANGE IN THE PRICES OF FACTORS OF PRODUCTION

A fall in the cost of pullets or of their food would reduce the cost of producing eggs. As a result more eggs could be supplied at the old price, or – looked at in another way – the original quantity could be produced at a lower price per egg. A rise in the wages of workers on chicken farms would have the opposite effect.

(4) CHANGES RESULTING FROM NATURE

(e.g. the weather, floods, drought, pest) and from *abnormal circumstances* (e.g. war, fire, political events).

(5) GOVERNMENT POLICY

A tax on the output of eggs or an increase in employers' national insurance contributions would result in fewer eggs being offered for sale at the old price. That is, the supply curve moves to the left. On the other hand a subsidy, by decreasing costs, would move the supply curve to the right (see p. 47).

Other changes in supply take longer, occurring through:

(6) IMPROVED TECHNIQUES

Technical improvements reduce costs of production, shifting the supply curve to the right. Thus improved automatic feeding devices might be developed, or selective breeding produce hens which lay more eggs.

(7) THE DISCOVERY OF NEW SOURCES OF RAW MATERIALS, OR THE EXHAUSTION OF EXISTING SOURCES

(8) THE ENTRY OF NEW FIRMS INTO THE INDUSTRY

3.3   The determination of price: market clearing

The demand and supply curves can be combined in a single diagram (Figure 3.3).

Let us see how this analysis helps as a first approach to understanding how the market is cleared. The assumptions we have made so far are:

(i) Many buyers and sellers.

(ii) Keen competition between buyers, between sellers, and between buyers and sellers.
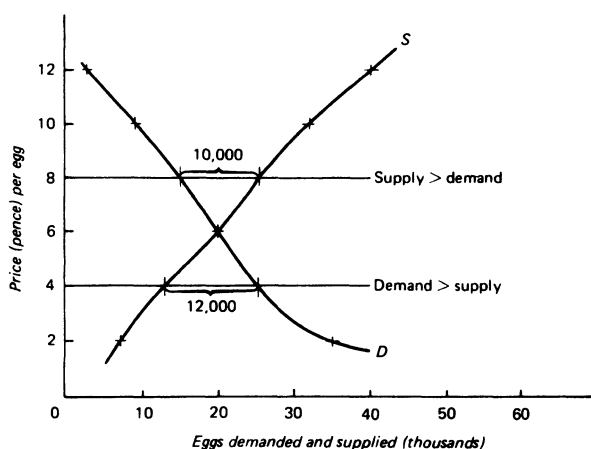
Figure 3.3 The determination of equilibrium price

(iii) More will be demanded at a lower price than at a higher price.
(iv) Less will be supplied at a lower price than at a higher.

Given assumptions (iii) and (iv), the two curves slope in opposite directions. Thus they cut at a single point – in our example, where the price is 6p. It can be predicted that in No-Such Market, where these conditions of demand and supply exist, the price of eggs will move towards and eventually settle at 6p. This is the *market clearing* or *equilibrium* price.

This proposition can be proved as follows. Suppose that initially the price of eggs is fixed at 8p. Here 15,000 will be demanded but 25,000 supplied. There is thus an excess supply of 10,000. But some sellers will want to get rid of their surplus supplies, and therefore reduce the price being asked. As this happens some supplies are withdrawn from a market, and there is an extension of demand. This continues until a price of 6p is reached, when 20,000 eggs are both demanded and offered for sale. Thus 6p is the only price at which there is harmony between buyers and sellers: given existing demand and supply, the market is 'cleared'.

Similarly, if the initial price is 4p, 25,000 will be demanded, but only 13,000 offered for sale. Housewives queue to buy eggs, and sellers see that their supplies will quickly run out. Competition among buyers will force up the price. As this happens, more eggs are supplied to the market, and there is a contraction of demand. This continues until a price of 6p is reached, when demand equals supply at 20,000 eggs.

3.4 Changes in the conditions of demand and supply

The equilibrium price will persist until there is a change in the conditions of either demand or supply.

Let us begin with our market price of 6p. Suppose tastes alter, and people eat more eggs. The conditions of demand have now changed, and the demand curve shifts to the right from D to D_1 (Figure 3.4).
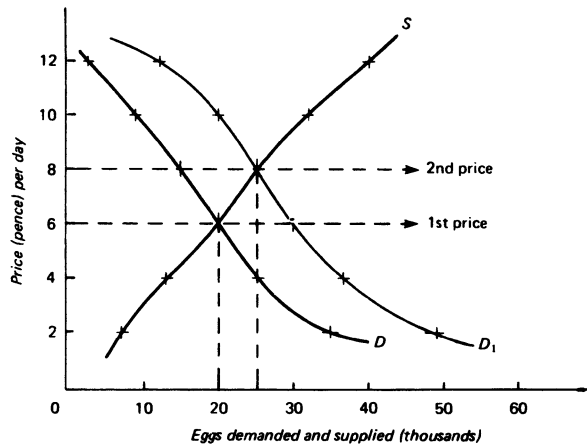
Figure 3.4 The effect on price of a change in the conditions of demand

At the original price of 6p we now have an excess of demand over supply – 30,000 eggs are demanded, but only 20,000 are supplied. As explained in the previous section, competition between buyers will now force up the price to 8p, where 25,000 eggs are both demanded and supplied.

Similarly a decrease in demand – resulting, for instance, from a significant fall in the price of tomatoes – would cause the curve to shift to the left and the price of eggs to fall.

Alternatively a change may occur in the conditions of supply. At any given price more eggs can be produced during the spring, when the supply curve shifts to the right from S to S_1 (Figure 3.5).
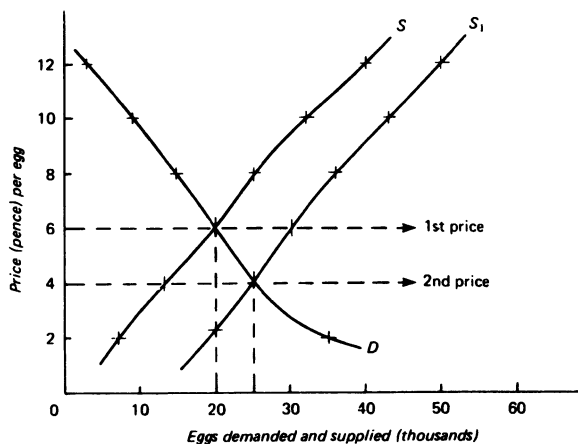
Figure 3.5 The effect on price of a change in the conditions of supply

At the original price of 6p we now have an excess of supply over demand – 30,000 eggs are supplied, but only 20,000 are demanded. Here competition amongst sellers will mean that the price falls to 4p, where 25,000 eggs are both demanded and supplied.

3.5   Functions of price in the market economy

Our analysis can be applied to practical problems, especially those relating to government policy. First, however, we use it to examine the role of price in the market economy.

In a free market, price both indicates and motivates.

(a) Price 'rations out' scarce goods

At any one time the supply of a good is relatively fixed. It therefore has to be apportioned among the many people wanting it. This is done by adjusting price. As price rises, demand contracts; as it falls, demand expands. At the equilibrium price demand just equals the supply. Should supply increase, the total quantity can still be disposed of by lowering the price; should supply decrease, the price would have to be raised.

We can illustrate how price works by considering two current problems:

(i) *Who shall be allowed to park his car in a congested area?*

Car parking is causing traffic congestion in the centre of Barthem City. This is because it costs motorists nothing to park their cars at the kerbside. The council decides to restrict parking to one side of the road and to 800 one-hour only places, each with a parking meter. The demand schedule for one-hour parking is estimated to be as follows:

Price (pence)	Demand
60	450
40	800
20	1 200
0	1 800

The council therefore fixes a charge of 40p.

This introduction of parking meters: (1) makes parkers pay for the space they occupy; (2) bars the all-day parker by limiting meter time to one hour; (3) forces the all-day parker and those who will not pay 40p to travel by public transport or to park off the street or out of the city centre; (4) causes the demand for off-street parking to increase, thereby encouraging firms to expand supply; (5) helps to relieve congestion by limiting parking to one side of the street or prohibiting it in busy roads.

(ii) *Why do ticket touts obtain such high prices for Cup Final tickets?*

To ensure that the regular football supporter can afford a Cup Final ticket, prices are fixed by the Football Association. Let us simplify by assuming that the FA has one price, £10 for the 100,000 tickets, but that a free-market price would be
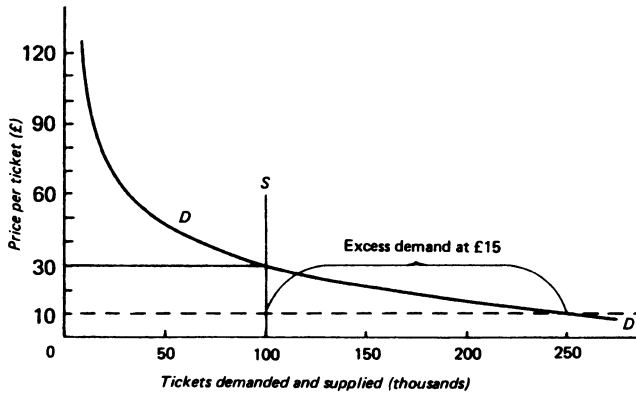
Figure 3.6   Excess demand for Cup Final tickets

£30. In Figure 3.6, when the price is £30 demand equals the available supply, but at the controlled price of £10 demand exceeds supply by 150,000.

But some tickets are obtained by touts, who re-sell at a profit in a free market where demand and supply determine price. Keen club supporters, not lucky enough to have been allocated a ticket, are willing to pay more than £10. As the price rises, some people possessing tickets may be induced to sell them to the touts. Thus the demand and supply curves are roughly as shown in Figure 3.7, giving a 'spiv-market' price of £100.

An important conclusion can be drawn from this example: where price is controlled below the market price, only some form of rationing can ensure that everybody gets a share of the limited supply. Normally this is achieved by the FA, which, after allocating a certain number of tickets to each finalist, limits each affiliated club to approximately two. One alternative would simply be a 'first
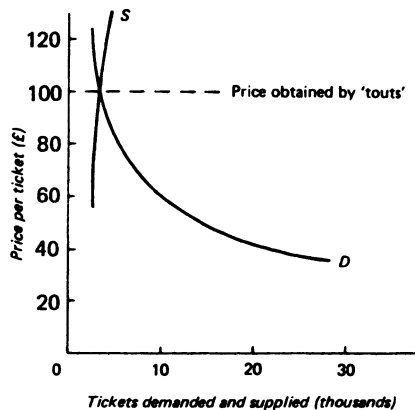
Figure 3.7   The black-market price of Cup Final tickets

come, first served' method of distribution, penalising those who could not queue
and increasing the scope for tout activity.

(b) Price indicates changes in wants

Prices are the signals by which households indicate the extent to which different
goods are wanted, and any changes in those wants.

Consider how the demand for owner-occupied houses in south-east England
has increased over the long-term through the pressure of population, higher real
income tax concessions, etc. As a result, prices have risen from OP to OP_1
(Figure 3.8).
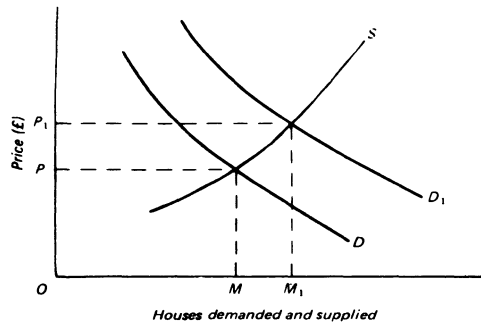
**Figure 3.8   The effect on house prices of an increase in demand by owner-
occupiers**

(c) Price induces supply to respond to changes in demand

When demand increases, price rises and supply extends; when demand
decreases, price falls and supply contracts. Thus in Figure 3.8 the increase in
price has made it profitable for extra houses MM_1 to be supplied by new building,
by transferring houses from the rented sector, etc.

(d) Price indicates changes in the conditions upon which goods can be
supplied

If the cost of producing a given commodity rises, this should be signalled to con-
sumers, who can then decide to what extent they are prepared to pay these higher
costs by going without other goods. Again this is achieved through price. Assume
in Figure 3.9 that the cost of producing good X has increased because raw materi-
als have risen in price. Where demand is depicted by D, most consumers pay the
higher costs (price rises by PP_1) rather than do without the good. Where demand
is depicted by D_1, consumers tend to go without the good as its price rises
(demand falls by MM_1), substituting other goods for it.
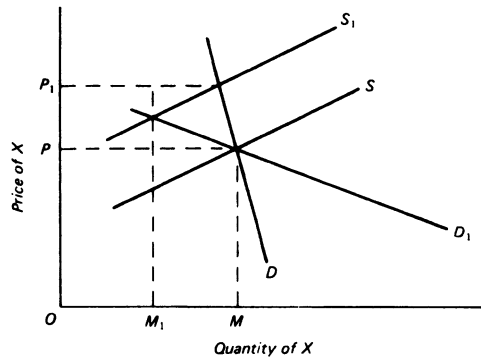
Figure 3.9   **The effect of a change in the conditions of supply on price and quantity traded**

(e) Price rewards the factors of production

Payments for factors of production give their owners spending power. The relative size of this spending power determines the division (usually termed 'distribution') of the cake produced. If the price of a good rises, producers can afford to offer higher rewards in order to attract factors from other uses.

3.6   Further applications

(a) How can the government stabilise commodity prices?

The government can use a stockpile in order to stabilise the price of basic commodities where demand or supply fluctuates.

In Figure 3.10, it is assumed that the demand for butter remains constant, but that the conditions of supply change from one period to another. S_1 is the supply curve for period 1, S_2 for period 2 and S_3 for period 3. The government has a stockpile by means of which it stabilises the price of butter at OP a tonne. This
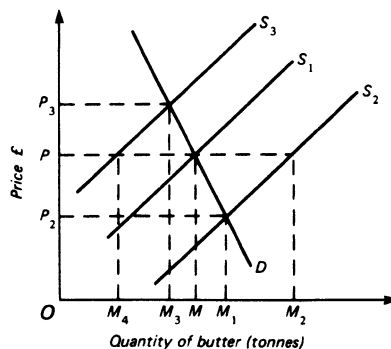
Figure 3.10   **Stabilisation of the price of butter**

it does by adding MM_2 to the stockpile in period 2 and withdrawing M_4M in period 3.

(b) How would an increase in the demand for cars affect the price of tyres?
Cars and tyres are 'jointly demanded'. With such goods prices move in the same direction. This can be seen in Figure 3.11. The increased demand for cars leads to an increased demand for tyres, and the prices of both rise.
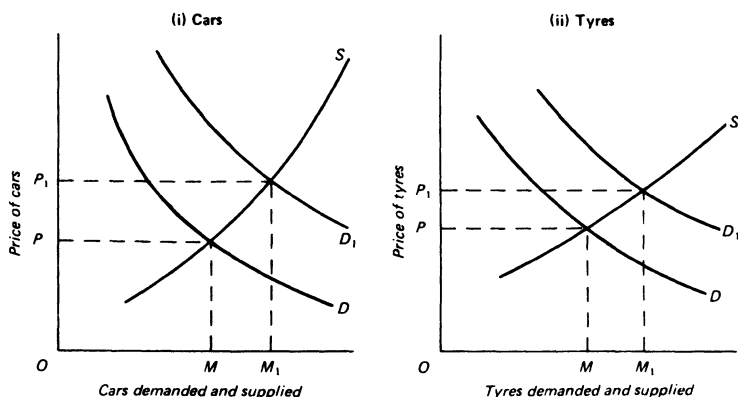
Figure 3.11   Joint demand

(c) How would an increase in the price of petrol affect the price of paraffin?
Petrol and paraffin are 'jointly supplied': increased production of one automatically increases production of the other. Suppose that demand for petrol increases but that there is no change in the demand for paraffin. The price of petrol rises from OP to OP_1, and supply expands from OM to OM_1 (Fig 3.12). But this means
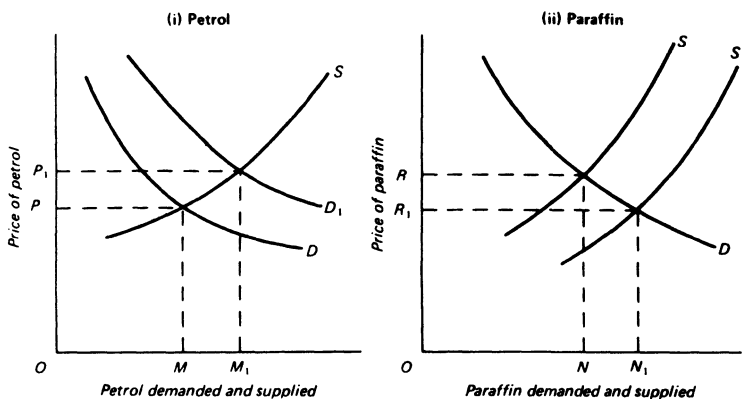
Figure 3.12   Joint supply

that the supply of paraffin increases, although there has been no change in price. Thus the supply curve for paraffin moves from S to S_1, and the price of paraffin falls from OR to OR_1.

(d) How could the government secure greater use of unleaded petrol?

Here the government must operate to alter the relative prices of leaded and unleaded petrol by increasing the tax on the former and reducing it on the latter. The effect is shown in Figure 3.13. Because of the increased tax on leaded petrol, demand decreases at all prices, the demand curve now being D_1 instead of D. Price, including the higher tax rises from OP to OP_1, and the amount sold contracts by MM_1 (Fig. (i)).
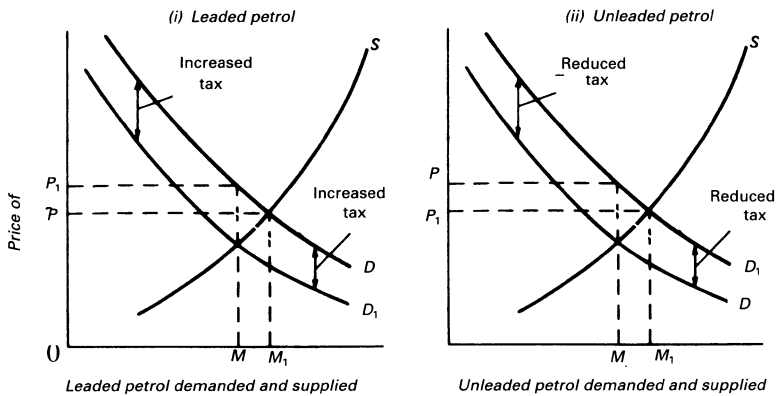
Figure 3.13   The effect on quantity bought of a change in the tax on a good

In contrast, the reduced tax on unleaded petrol allows the same amount to be obtained at a lower price, as shown by the higher demand curve, D_1. Price falls to OP_1 and the amount sold expands by MM_1 (Fig. (ii)).

④ **Demand**

4.1 The marginal-utility theory

Our assumption in Chapter 3 that more of a good will be demanded the lower its price was based solely on our everyday observations. However, by examining a little more closely how the individual consumer 'economises', we can explain why this is normally so. We shall use the marginal-utility theory.

(a) Preliminary assumptions
  (i) Our consumer is a housewife.
  (ii) She has a limited housekeeping allowance per week.
 (iii) She acts so as to obtain the maximum satisfaction from her limited income.
 (iv) During the period of time under consideration, income, tastes and the other conditions of demand do not change.
  (v) She knows how much satisfaction each unit of a good will give.
 (vi) She is one of a large number of buyers and her demand does not directly affect the price of the good.

(b) Questions to be answered
Three basic questions have to be answered. First, what conditions will hold when the consumer has obtained the maximum satisfaction from her limited resources? In other words, what are the equilibrium conditions? Second, how does she achieve this equilibrium? Third, what happens when the equilibrium is disturbed by a price change? We deal with each in turn.

(i) *The equilibrium condition*
Our housewife will be in equilibrium when she would not switch a single penny of her expenditure on one good to another.

    We can be more explicit by introducing the term 'utility'. In economics this simply means that a good has the power to satisfy a want, irrespective of whether it is useful. Note, too, that we cannot measure utility objectively; like fear, it is purely subjective to the individual.

    However, our housewife knows in her own mind how much satisfaction each good affords her. She is in equilibrium, therefore, when she has obtained the greatest possible utility from her income: that is, *she maximises total utility.*

    She achieves this by careful allocation of her spending – say between cheese and margarine. All the time she is asking: 'If I spend a penny more on cheese, will I obtain more or less utility than if I spent the penny on margarine?' Only

when the satisfaction she obtains from the last penny spent on cheese (in the sense of the penny she only just decided to spend) is equal to that from the last penny spent on margarine will she be in equilibrium. That is, her spending adjustments take place at the margin.

Note that we did *not* say that she obtained the same utility from the last pound of cheese as she obtained from the last pound of margarine. If, for instance, a pound of cheese were four times as expensive as a pound of margarine, that would obviously be unreasonable; we would expect four times the amount of utility.

Sometimes, however, we cannot buy goods in 'pennyworths' – the good is 'lumpy' and we have to take a whole 'lump' of it or nothing at all. Can we re-state our equilibrium condition to allow for this? Yes, but first we must define more carefully the concept of the margin and what we mean by 'marginal utility'.

Each small addition to a given supply of a good is called the *marginal incre-ment,* and the utility derived from this increment is known as the *marginal utility.* Our original condition of equilibrium can therefore be stated in general terms:

$$\text{The marginal utility of 1p spent on good } A = \text{The marginal utility of 1p spent on good } B$$

But the marginal utility of 1p spent on good A depends on how much of a unit of good A you get for 1p. Thus:

$$\text{The marginal utility of 1p spent on good } A = \frac{\text{The marginal utility of one unit of good } A}{\text{The number of pence it costs to buy a unit of good } A}$$

Similarly with good B. Thus our original equilibrium condition can be rewritten as:

$$\frac{\text{The marginal utility of one unit of good } A}{\text{Price of a unit of } A \text{ in pence}} = \frac{\text{The marginal utility of one unit of good } B}{\text{Price of a unit of } B \text{ in pence}}$$

That is:

$$\frac{\text{Marginal utility of good } A}{\text{Price of } A} = \frac{\text{Marginal utility of good } B}{\text{Price of } B}$$

The argument can be extended to cover more than two goods.

(ii) *How does the consumer achieve this equilibrium?*
The question must now be asked: How can our housewife arrange that the utility of the last penny spent on different goods is the same? The answer is to be found
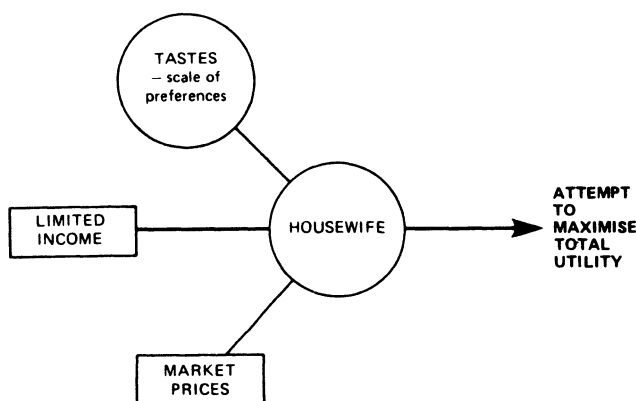
Figure 4.1   Factors affecting the equilibrium of the housewife

in the so-called *law of diminishing marginal utility*. Although wants vary con-
siderably in their nature, they all possess the underlying characteristic that in a
given period they can be satisfied fairly quickly. Thus, if a boy drinks lemonade
to quench his thirst, the first glass will yield him a great amount of satisfaction.
Indeed, the second glass may be equally satisfying. But it is doubtful whether
he will relish the third glass to the same extent. If he continues to drink the lem-
onade, there will come a time when a glass gives him no additional satisfaction
whatsoever and, in fact, it might be that he would be better off without it –
there is a 'disutility'. We can therefore state a general rule that the utility
derived from any given addition to a consumer's stock of a good will
eventually decline.

This means that our housewife can arrange that equal utility is derived from
the last penny spent on each good by varying the quantity she buys. If she buys
more of a good, the stock of other goods remaining fixed, its marginal utility
relative to other goods falls. Similarly, if she reduces the quantity she buys,
the marginal utility of the good relative to other goods rises. She makes such
marginal adjustments until she is in equilibrium.

(iii) *What happens when the equilibrium is disturbed by a price change?*
Suppose the price of cheese falls from 130p to 110p per pound while the prices of
other goods remain unchanged. How will this affect her demand for cheese? We
can proceed in either of two ways:

(1) The fall in the price of cheese will enable her to obtain more cheese than
before for every penny, including the last, which she was spending on it. More
cheese usually implies greater satisfaction. The last penny spent on cheese, there-
fore, now yields greater satisfaction than the last penny being spent on other
goods. Hence she reduces the utility obtained from the last penny spent on cheese
by buying more cheese.

(2) The alternative form of the equilibrium condition is:

$$\frac{\text{The marginal utility of the last lb of cheese}}{\text{Price of lb of cheese}} = \frac{\text{The marginal utility of one unit of good } B}{\text{Price of one unit of good } B}$$

A fall in the price of cheese destroys this relationship; the marginal utility of cheese to its price is now higher than with goods B, C, etc. To restore the equilibrium relationship, the marginal utility of cheese must be decreased. Hence our housewife buys more cheese.

The reasons for this expansion in the demand for cheese can be analysed more closely. A reduction in the price of cheese means that our housewife is now able to purchase all the cheese she had before and still have money left over. This is an *income* effect of a price fall – she can now buy more of all goods, not only of cheese. But, in addition to this income effect, more cheese will tend to be bought because of a 'substitution effect'. At the margin this means that a penny spent on cheese will now yield more satisfaction than a penny spent on other foods. Thus cheese is substituted for other foods. If cheese is a good substitute, marginal utility will diminish comparatively slowly as the consumption of it increases. A given price fall, therefore, will lead to a considerable increase in the quantity of cheese demanded.

Although we have explained the behaviour of only one consumer, it is reasonable to expect other buyers in the market to act similarly. Since the *market-demand curve* is made up of the demand schedules of all the individual purchasers, we can conclude that more of a good will be demanded the lower its price.

4.2 Price elasticity of demand

(a) Measurement of elasticity of demand

Consider Figure 4.2. At price OP, demand for both commodities A and B is OM. But when the price of both falls by PP_1, demand for A expands by only MM_1, whereas that for B expands by MM_2. Responsiveness of demand to a change in price is of obvious importance to a firm which has some control over the price
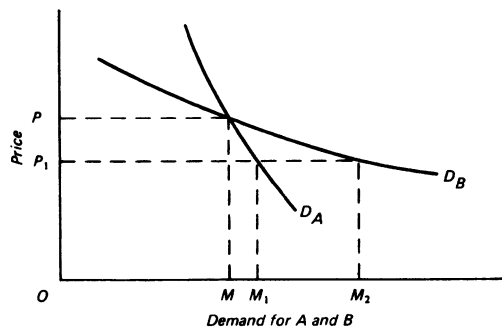
Figure 4.2   Elasticity of demand

it charges. In economic terms it is interested in its 'price-elasticity of demand'.

Elasticity of demand always refers to the elasticity at a particular price, and in what follows when we talk about 'elasticity', it will be assumed that there is some price in mind.

Elasticity of demand is defined by comparing the rate at which demand expands to the rate at which price falls. Using this definition, elasticity of demand can be measured in two ways. One is direct, showing the degree of elasticity: the other is indirect, merely indicating whether the demand for the good is elastic or inelastic.

(i) Elasticity of demand is the proportionate change in the amount demanded in response to a small change in price divided by the proportionate change in price. That is:

$$\text{Elasticity of demand} = \frac{\text{Proportionate change in demand}}{\text{Proportionate change in price}}$$

$$= \frac{\dfrac{\text{Change in quantity demanded}}{\text{Original quantity demanded}}}{\dfrac{\text{Change in price}}{\text{Original price}}}$$

$$= \frac{\dfrac{\text{New quantity} - \text{Old quantity}}{\text{Old quantity}}}{\dfrac{\text{New price} - \text{Old price}}{\text{Old price}}}$$

We can illustrate by an example from the demand schedule, Table 3.1 (see p. 34). When price falls from 10p to 8p, demand for eggs expands from 9,000 to 15,000. Elasticity of demand is thus equal to

$$\frac{\dfrac{6,000}{9,000}}{\dfrac{2}{10}} = \frac{\dfrac{2}{3}}{\dfrac{1}{5}} = 3\tfrac{1}{3}$$

Similarly, for a rise in price from 8p to 10p, elasticity of demand equals $1\tfrac{3}{8}$. The difference in the two results occurs because we are measuring from different prices and for a relatively large change. (It should also be noted that it is conventional to ignore the minus sign which results from the fact that the typical demand curve has a negative slope.)

(ii) If the proportionate expansion in demand is greater than the proportionate fall in price, the total amount spent on the good will increase. In other words demand is elastic when, in response to a fall in price, total outlay increases; or, in

response to a rise in price, total outlay decreases. Similarly demand is inelastic when, in response to a fall in price, total outlay decreases; or, in response to a rise in price, total outlay increases. Elasticity of demand is equal to unity when, as price changes, total outlay remains the same. Thus, using the same demand schedule, we have:

Table 4.1 Elasticity of demand and total outlay

Price of eggs (pence)	Demand (thousands)	Total outlay (pence)	
10	9	90 000	} Elastic demand
8	15	120 000	
6	20	120 000	} Inelastic demand
4	25	100 000	

Between 8p and 6p, elasticity of demand equals unity.

(b) Factors determining elasticity of demand

(i) *The availability of substitutes at the ruling market price*
As a good falls in price, it becomes cheaper relative to other goods. People are induced to buy more of it to replace goods which are now relatively dearer. How far they can carry out this replacement will depend upon the extent to which the good in question is, in their own minds, a substitute for the other goods. Goods within a particular class are easily substituted for one another. Beef is a substitute for lamb. Thus, if the price of beef falls, people will buy more beef and less lamb. Between one class and another, however, substitution is more difficult. If the price of meat in general falls, there will be a slight tendency to buy more meat and less fish, but this tendency will be very limited because meat is not nearly so perfect a substitute for fish as beef is for lamb.

The success of supermarkets has been based on the high elasticity of demand for their products; people switch to them when prices of processed goods are reduced, for they can recognise the packages and tins as being almost perfect substitutes for those being sold at higher prices by other retailers.

(ii) *The number of possible substitute uses*
Where a good can be substituted for another good, its demand tends to be elastic. And the more goods it can be substituted for, so the more will demand for it extend as its price falls. Thus reductions in the price of plastics have led to large extensions of demand as they have been substituted for materials used in such articles as enamel bowls, galvanised buckets, paper wrappings, glass garden cloches, wooden toys and tin containers.

(iii) *The proportion of income spent on a good*
When only a very small proportion of a person's income is spent on a good, as for example with pepper, salt, shoe-polish, newspapers and toothpaste, no great effort is made to look for substitutes when its price rises. Demand for such goods, therefore, is relatively inelastic.

On the other hand, when the expenditure on a good is fairly large – as, for example, with meat – a rise in price would provide considerable incentive to find substitutes.

(iv) *The period of time*
Since it takes time to find substitutes or to change spending habits, elasticity may be greater the longer the period of time under review. In practice many firms try to overcome the ignorance or conservatism of consumers by advertising, giving free samples or making special offers.

(v) *The possibility of new purchasers*
In discussing the possibility of substitution above, we have looked at elasticity of demand solely from the point of view of the individual consumer. But when we are considering market demand we must allow for the fact that, as price falls, new consumers will be induced to buy the good. In fact, with goods such as cars, video recorders, washing-machines, etc., of which most people require only one, it is the fall in price bringing the good within reach of new consumers which leads to the increase in demand. Hence a fall in price which induces people in a numerous income group to buy will produce a big elasticity of demand.

(c) Uses of the concept of elasticity of demand
The concept of the elasticity of demand figures prominently in both the economist's theoretical analysis (see p. 93) and in the practical decisions of the businessman and government.

Thus a trade union will find it more difficult to obtain a wage increase for its members without creating unemployment where the elasticity of demand for the product made is high (see p. 143).

British Rail, too, have to consider elasticity of demand when fixing fares. Should they, for example, raise fares in order to reduce losses? If, at existing fares, the demand is relatively elastic, then a fare increase would mean that total revenue would fall. Losses would only be reduced if operating costs (through carrying fewer passengers) fell more than revenue. Indeed, consideration could be given to reducing fares, since the extra revenue might cover any additional cost of running more trains.

Finally, the Chancellor of the Exchequer must take account of elasticity of demand when imposing a selective tax on a particular good. The demand may be so elastic that the increase in price might cause such a falling-off in sales that the total tax received is less than it originally was.

4.3 Income-elasticity of demand

An increase in real income usually increases the demand for goods, but to a vary-ing degree. Thus it is possible to speak of *income-elasticity of demand* – the proportionate change in demand divided by the proportionate change in real income which has brought it about. If demand increases 20%, for instance, as a result of a 10% increase in real income, income-elasticity of demand equals 2. Which goods have a high income-elasticity of demand depends upon current living standards. In Western Europe today it is demand for such goods as cars, dishwashers, central-heating appliances and personal services which expands most as income increases. In contrast, necessities, such as salt and soap, have a low income-elasticity of demand.

4.4 Cross-elasticity of demand

Where two goods are related, e.g. as substitutes or complements, a change in the price of one will lead to a change in demand for the other. Thus a rise in the price of oil leads to an increase in the demand for coal, while a fall in the price of video recorders leads to an increased demand for video tapes.

The extent to which the demand for a good changes in response to a price change of another good is known as *cross-elasticity of demand:*

$$\text{Cross-elasticity of demand} = \frac{\text{Percentage change in the quantity demanded of good } X}{\text{Percentage change in the price of good } Y}$$

With substitutes, cross-elasticity is positive. For example, an increase in the price of Y would lead to an increase in the demand for X (as with oil and coal in the example above). With complements, cross-elasticity is negative, since a fall in the price of Y leads to a rise in the demand for X (as with video recorders and tapes above). The closer the substitutes or complements, the larger will be the figure for cross-elasticity. A cross-elasticity near zero signifies that there is little relationship between the two goods (Figure 4.3).

Cross-elasticity

Complementarity $\longleftarrow$ $\ominus$ $\xrightarrow{+}$ Substitutability
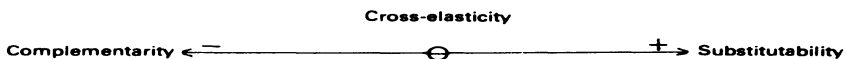
Figure 4.3   Cross-elasticity of demand

⬡ 5 ⬡ Supply: (i) the structure of industry

5.1 The role of the firm

(a) The meaning of 'production'

Early economists considered that only work in the extractive industries (agriculture, mining and fishing) was productive. In his *Wealth of Nations* in 1776 Adam Smith added manufacturing, but he was specific in excluding workers who merely rendered services.

This was illogical. People work, and production takes place, in order to satisfy wants. Consequently people who render services must be regarded as being productive. The soldier, actor and footballer are all satisfying wants. Similarly, in a factory, the clerk who calculates the wages is just as productive as the man who makes the nuts and bolts. All are helping to produce the final product, a good satisfying wants.

Wants can take different forms. Most people like a newspaper to read at the breakfast-table; thus the boy who takes it from the shop to the customer's letter-box is productive. Most people, too, prefer to buy their potatoes weekly; thus the farmer or merchant who stores them through the winter is satisfying the wants of consumers, and is similarly productive. Utility is created by changing not only the *form* of our scarce resources, but also their *place* and *time*.

For certain purposes it may still be useful to classify industries broadly. *Primary industries* cover the first steps in the productive process – agriculture, fishing, mining and oil-prospecting. *Secondary industries* use the raw materials of the extractive industries to manufacture their own products – flour, clothing, tinned salmon, steel girders, petrol and so on. *Tertiary industries* are concerned with the provision of services – transport, communications, distribution, commerce, government, and professional and other services.

(b) The objectives of the firm

In a market economy a firm has to cover its costs if it is to stay in business. Thus regard must be paid to 'profitability'.

But in practice are firms always single-minded in seeking to *maximise money profits*? The answer is no; there is a range of possible objectives.

Personal motives may be important, especially where the manager is also the owner of the firm. Thus emphasis may be placed on good labour relations, the welfare of the workers, the desire for power, political influence, public esteem or simply 'a quiet life'. To cover such objectives profit would have to be interpreted in a wider sense than 'money profit'.

With major companies there is in practice a gap between the ownership and administration. The business is run by professional managers, and is too complex for shareholders to be able to exert effective control. This applies even to the institutional shareholders, who avoid being directly involved in the running of the business. Thus the motives of the full-time executive managers tend to over-ride the shareholders' desire for maximum return on capital invested. Managers may be anxious for the security of their own jobs and, instead of taking the calculated risks necessary to earn maximum profits, may tend to play for safety. More likely, they will be motivated by personal desires for status. Provided they achieve a level of profit which keeps shareholders content, their positions and salaries can be enhanced by expanding the firm to where it *maximises sales* rather than profits. Alternatively, the rate of growth may be maximised.

Even when there is an emphasis on money profit, a firm may stress its long-term position rather than immediate maximum profit. Security of future profits may be the dominating motive for mergers and takeovers as an alternative to developing new products and techniques. Moreover, where there is an element of monopoly, a firm can follow its own pricing policy rather than have it determined by competitive market conditions (see Chapter 6). In such circumstances it may not adjust prices to short-term changes in demand and supply conditions. For one thing, there are the administrative costs of printing and distributing new price lists. For another, frequent changes in price tend to offend retailers and customers.

Again, a firm enjoying a degree of monopoly has always to assess what effect the pursuit of maximum profit may have on its overall position in the long term. Will a high price attract new entrants or encourage the development of a rival product? Will it lead to adverse publicity and eventually to government intervention by a reference to the Office of Fair Trading?

Finally, a firm has often to modify its objectives in deference to government policy. Thus it may be expected to follow government guidelines regarding wage increases, to have regard to the environment in the disposal of its waste products and even to retain surplus workers for a time rather than add to an already high level of unemployment.

Yet, while we must take account of these other objectives, our analysis cannot proceed far if any are seen as the main motive force of the firm. In any case they merely supplement the profit objective, for profits have to be made if the firm is to survive. Thus it is useful to start with the broad assumption that firms seek to maximise profits. We can then establish principles concerning how resources should be combined and what output should be produced.

(c) The decisions of the firm

To achieve its objective of maximising profit, a firm has to assess the demand of potential customers for its product and produce that output which secures the greatest difference between total revenue and total cost. Moreover, the cost of producing this given output must be the lowest possible.

This means that the firm has to answer the following questions:

(i) What goods shall it produce?
(ii) What shall be its legal form?

   (iii) How shall it raise the necessary capital?
   (iv) What techniques shall be adopted, and what shall be the scale of operations?
   (v) Where shall production be located?
   (vi) How shall goods be distributed to the consumer?
  (vii) How shall resources be combined?
 (viii) What shall be the size of output?
   (ix) How shall it deal with its employees?

We consider the first six problems in the remainder of this chapter; the rest are examined in Chapters 6 and 10.

5.2   What to produce

(a) The first approach

Other things being equal, a firm will produce those goods which enable it to make the greatest return on capital. However, in practice, this usually means that it has to choose a line of production within the limited range of its specialist knowledge. Let us assume that the firm is manufacturing light farm machines and that it is contemplating producing lawnmowers.

Since it is likely that some firms are already producing lawnmowers, the market economy throws up two guidelines. First, there is the current price of mowers. The firm would have to estimate its own costs for producing similar mowers, the number it could expect to sell at this price, and its likely profits, and thus calculate the return on capital employed. Second, the accounts of companies have to be filed with the Registrar of Companies, and the profit earned by public companies is publicised in the financial pages of leading newspapers and specialist journals. If existing producers of lawnmowers were shown to be earning a high rate of profit, the prospects for a new competitor would look favourable.

(b) Market research

Where the proposed market is new or different from that for existing products, the above indicators are not so useful. Here the firm must fall back on some form of market research.

Initially, it may be producing similar goods, e.g. light agricultural machinery, and some indications of potential demand may come from wholesalers, retailers or even customers in conversation with the firm's representatives. Such suggestions can be cross-checked with those of other distributors .

Where the reaction is generally favourable, more thorough market research can be carried out, probably through a specialist market-research organisation. Market research can cover desk research, field studies and test marketing.

Desk research examines the broad determinants of the potential demand by using (i) published material, e.g. government statistics, and (ii) the firm's own sales records. As we saw in Chapter 3, these determinants are price and the various conditions of demand. More specialist facts could be obtained from relevant periodicals and trade journals, e.g. *Gardeners' World* (where circulation figures

indicate the number of keen gardeners). Membership figures for the Royal Horticultural Society could also be used.

More precise information on potential sales necessitates a planned consumer-orientated *market-research programme* in potential markets. This would cover many aspects of market behaviour, particularly consumer reaction to the product – especially with regard to its quality, packaging, delivery dates and after-sales service, and to price cuts.

Before a national or major sales campaign is undertaken, some form of *test marketing* would probably be carried out so that modifications could be made to correct any deficiencies. For instance, such a test might reveal that certain features of the product were unnecessary, thus permitting greater standardisation. Moreover, not all potential customers have identical preferences. The firm would therefore consider (i) a 'marketing mix' – producing different models at different prices – and (ii) varying sales methods and channels of distribution (see below).

5.3   The legal form of the firm

In the private sector a firm can trade as a sole proprietor, partnership, private company, public company or co-operative society. For a new firm the choice is really between the first three, the actual decision largely resting on whether freedom from control by the Registrar of Companies compensates for unlimited liability. The legal form hardly affects its ability to raise capital for, unless it is an offshoot of a large parent company, it has to be fairly successful before outsiders can be induced to subscribe capital for large-scale development.

(i) *The sole proprietor*
The sole proprietor is the oldest form of business organisation. Even today, from the point of view of numbers, small firms predominate, but in their total productive capacity they are far less important than companies (see Table 5.1, p. 72). Such one-person firms range from the window-cleaner working on his own account to the farmer, shopkeeper and builder who employ other workers and may even own many separate units. Nevertheless these businesses all share the characteristic of being owned and controlled by a single person. This person decides the policy of the firm, and alone takes the profits or bears losses. This makes for energy, efficiency and careful attention to detail. In addition, the only accounts which have to be submitted are to the Board of Inland Revenue for income-tax assessment and to the Customs and Excise Department if registered for VAT. No corporation tax is payable.

However, the sole proprietor suffers from five main disadvantages. First, such a firm can only develop slowly, because sources of capital are limited. The success of the venture, especially in its early stages, depends mainly on the person in charge, and nobody is likely to provide capital unless there is that confidence in the proprietor which comes from personal contact. Hence the main source of capital is the owner's savings together with such additional sums as can be borrowed from relatives, close friends, a bank, or perhaps the Rural Development Commission. In time, development may be financed by ploughing back profits,

but this will probably be a slow process and sole traders generally remain comparatively small.

Second, in the event of failure, not only the assets of the business, but also the private assets and property of the proprietor can be claimed by creditors. In short, there is no limited liability.

Third, where profits are high, income tax paid on annual profits may be larger than a company's corporation tax. This is because income may be taxed at a high marginal rate of tax, whereas corporation tax is only 25% if a company's profits are less than £250,000 (with some marginal relief up to £1,250,000). Less tax leaves more funds for investing in the business.

Fourth, it is more difficult to transfer part of a business than to transfer shares in a company.

Fifth, there is a lack of continuity; on the retirement or death of the owner, a one-man firm may cease to function.

Because of these disadvantages, sole proprietors are mainly confined to businesses which are just setting up and also to certain industries, such as agriculture and retailing, where requirements of management make the small technical unit desirable.

(ii) *The partnership*

More capital is available when persons join together in a 'partnership', though normally not more than twenty may do so. Each partner provides a part of the capital and shares the profits on an agreed basis. Yet the amount of capital which can be raised in this way is still inadequate for large-scale business. Thus partnerships remain relatively small, predominating in retailing, insurance broking and underwriting, and among professional people (doctors, surveyors, consulting engineers and lawyers), where the capital provided is not so much in the form of money as in experience and skill, each partner probably specialising in a particular branch.

Nor is the partnership without its snags. The risk inherent in unlimited liability is increased because all partners are liable for the firm's debts, irrespective of the amount of capital which each has invested. Only if a partner takes no share in the management of the firm, and there is at least one ordinary partner, can he enjoy limited liability. Second, since any action taken by one partner is legally binding on the others, not only must each partner have complete confidence in his fellows, but the risk inherent in unlimited liability increases with the number of partners. Finally, at any time one partner may give notice to end the partnership, and it is automatically dissolved upon the death or bankruptcy of a partner. To preserve the business, surviving partners may be put to great expense and trouble in buying the partner's share or finding a purchaser acceptable to everyone.

(iii) *The joint-stock company*

The joint-stock company dates from Tudor times, when England's foreign trade began to expand. Instead of a trading ship's being owned by one person, it was financed by a number of people who bought 'shares' in a company formed for the

purpose. However, since they enjoyed no limited liability, people were reluctant to join such companies: by purchasing only one share a person risked not only the money he had invested, but all his private assets, should the company be forced into liquidation. This made it impossible to adopt the technique of spreading risks by investing in a number of companies.

The industrial revolution, with the introduction of machines and factory organisation, made it essential that more capital should be available to industry. So, in order to induce small savers to invest, parliament granted limited liability in 1855.

Today the joint-stock company is the most important form of business organisation. The advantages it enjoys over the partnership are limited liability, continuity, the availability of capital (since investors can spread their risks and sell their shares easily) and, should the need arise, ease of expansion. Indeed, some kinds of businesses, e.g. aircraft production, could not be operated on a small scale. Here firms have to start as joint-stock companies, being either sponsored by important interests or developed as subsidiaries of existing large firms.

Against these advantages, however, the small joint-stock company in particular has to consider certain snags. Additional cost is incurred in submitting audited annual accounts and an Annual Return to the Registrar of Companies. Furthermore, any assets of the company which have been built up over the years will increase the value of the original shares (usually owned by a family), so that when the time comes to wind up the company, e.g. because of retirement, this increase may be subject to capital gains tax. Finally, if the company is expanded by the issue of more shares, the original owners may lose control or even be subject to a takeover bid.

Joint-stock companies are of two main kinds, private and public.

(1) THE PRIVATE COMPANY

A private company is simply a company that is not a public company. While conferring the advantage of limited liability, it allows a business to be privately owned and managed. The formalities involved in its formation are few, but the Companies Act 1948 imposes conditions restricting its size and the sale of shares to the public.

Thus the private company is particularly suitable for either a medium-sized commercial or industrial organisation not requiring finance from the public, or for a speculative venture where a small group of people wish to try out an idea and are prepared to back it financially to a definite limit before floating a public company. While private companies are far more numerous than public companies, their average capital is much smaller.

(2) THE PUBLIC COMPANY

To obtain a large amount of capital it is necessary to form a public company. This must have at least two shareholders, a minimum authorised capital of £50,000 and carry the designation 'public limited company' – abbreviated to PLC – after its name.

But it is the second step which is really important – getting its shares quoted on the Stock Exchange. This involves an exhaustive examination and full advertisement of the company's affairs. Sufficient shares have to be available to outsiders to make dealing, and the price fixed, realistic.

(iv) *Co-operative societies*

Although there were many co-operative societies in operation before the Rochdale Pioneers of 1844, it was these twenty-eight workers who started the modern co-operative movement. By subscribing a few pence per week they accumulated an initial capital of £28, with which they rented a store and started trading with small stocks of flour, oatmeal, sugar, butter and candles. Profits were distributed to members in proportion to their purchases. In 1991 there were 65 retail co-operative societies in the UK, with an aggregate membership of over 8 million. Turnover was £7.3 bn, making these societies Britain's biggest retailer. In addition, these retail societies largely provide the capital and control the operations of the Co-operative Wholesale Society.

The minimum shareholding in a retail co-operative society is usually £1. Only if a full share is held does a member enjoy voting rights, but each member has only one vote irrespective of the number of shares held. Some societies distribute profits as a dividend in proportion to the value of the member's purchases over the period as recorded at check-out through a numbered plastic card. Others use the National Dividend Stamp scheme run by the Co-operative Wholesale Society. Stamps are given to customers in proportion to their purchasers, and a book of stamps can be redeemed for cash, goods or a deposit in a share account. Both systems have allowed co-operative shops to compete with supermarkets.

Co-operative societies described above are organised directly by consumers and are therefore called 'consumer co-operative societies'. Producers have also formed 'producer co-operative societies' e.g. the Meriden motorcycle workers' co-operative, which was established with government aid when its firm was threatened with closure. A highly successful retail co-operative is the John Lewis Partnership; and building societies can be regarded as 'co-operative' ventures.

These co-operatives are chiefly important in agriculture, particularly where production is carried on by small farmers, as in Denmark, New Zealand and Spain. Their main function is to market their produce and to purchase inputs. They are likely to become more important in the UK as the government winds up the Marketing Boards.

5.4   Raising the necessary capital

(a) The need for liquid capital

In order to employ factors of production, a firm has to have finance. This is usually divided into (i) working capital and (ii) fixed capital.

(i) *Working capital* is for purchasing 'single-use' factors – labour, raw materials, petrol, stationery, fertilisers, etc. – more or less the factors we refer

to in chapter 6 as 'variable factors'. Finance for working capital can be obtained from a variety of sources: banks, trade credit, finance companies, factor houses, tax reserves, inter-company finance, advance deposits from customers and the government (e.g. through the Enterprise Allowance Scheme which on conditions provides an allowance of £40 per week for a year when starting a new business).

(ii) *Fixed capital* covers factors which are used many times – factories, machines, land, lorries, etc. Some finance for fixed capital is therefore required initially for advance payments on factory buildings, machinery and so on before the firm is earning revenue, though it is possible to convert fixed capital into working capital by renting buildings, hiring plant and vehicles or by leasing or buying on deferred payments through a finance company. Normally, fixed-capital requirements are larger than those for working capital. Moreover, lenders recognise that they part with their money for a longer period and accept a greater risk. Thus finance for fixed capital tends to be more difficult to raise than for working capital, unless the business starts as an off-shoot of a parent company.

(b) The long-term capital of a company
The *long-term finance* of a company is obtained in four main ways: (a) selling 'shares' in the company; (b) borrowing; (c) obtaining a government grant or loan; (d) retaining profits.

(i) *Shares*
A 'share' is exactly what the name implies – a participation in the provision of the capital. Shares may be issued in various units, usually from 5p upwards, purchasers deciding how many they want. Such an investment, however, involves two main risks. First, profits may be disappointing, and the price of the share may fall. Second, share prices in general may be falling just when the owner wishes to sell. To minimise these risks, investors usually have a portfolio of different shares, debentures and government bonds.

(1) ORDINARY SHARES
The dividend paid to the ordinary shareholder depends mainly upon the profitability of the company. However, the ordinary shareholder's dividend ranks last in order of priority, and if the company should be forced into liquidation the ordinary shareholder is repaid only after other creditors have been paid in full. Thus the ordinary share is termed 'risk capital'. In return each ordinary shareholder has a say in the running of the company, voting according to the number of shares held. At a general meeting directors are appointed or removed, changes made in the company's methods of raising capital and conducting business, and auditors appointed. Thus, because they take the major risks and decisions regarding the policy of the company, the ordinary shareholders are the real 'entrepreneurs'. In practice, however, their rights are rarely exercised. Usually few shareholders take the trouble to attend meetings, while unless the company is large the directors

may control a high proportion of the shares and so be in a strong position. Indeed this may be achieved by making all new shares 'non-voting "A" shares'. Thus directors tend to be self-perpetuating.

(2) PREFERENCE SHARES

If investors prefer a slightly reduced risk, they can buy preference shares. Here the dividend is fixed at a given percentage and paid before the ordinary share-holder. Should the company be forced into liquidation, the preference share-holder usually ranks above the ordinary shareholder when it comes to the redemption of capital. Preference shares may be 'cumulative': if the company cannot pay a dividend one year, arrears may be made up in succeeding years before ordinary shareholders receive any dividend.

Since 1965 preference shares have lost popularity because of their unfavourable tax treatment (see below).

(ii) *Borrowing*

Long-term loans are usually obtained by issuing 'debentures', redeemable after a specified period but bearing a fixed rate of interest. Since this is a first charge on the company's profit, the risk to the investor's income is not so high. Moreover, should the company fail, debenture-holders are paid first. In fact 'mortgage debentures' are secured on a definite asset of the company. Should the company be unable to meet its interest charges or to redeem the loan when due, the debenture-holders can force it into liquidation.

Thus the purchaser of a debenture takes less risk should the company fail. But because he is merely lending money he enjoys no ownership rights of voting on management and policy. On the other hand, a company whose profits are subject to frequent and violent fluctuations is not well-placed to raise capital through debentures. The method is best suited to a company making a stable profit (adequate to cover the interest payments) and possessing assets, such as land and buildings, which would show little depreciation were the company to go into liquidation.

A company having a large proportion of fixed-interest loans to ordinary shares is said to be 'highly geared'. Such a company will be able to pay high dividends when profits are good, but unable to make a distribution when profits are low. Where profits are expected to rise in the future, therefore, a company may prefer to raise capital for expansion by issuing debentures if the cost of doing so is not too high. But it is the present-day corporation tax which provides the main impulse in this direction. Debenture interest (but not preference-share interest) is an accepted cost for the purpose of calculating tax. Thus it reduces taxable profits. On the other hand, if finance is raised by shares, there is no prior interest charge, and profits (which are subject to tax) are higher by that amount. This tax advantage has, since the introduction of corporation tax in 1965, led companies to finance capital expansion as far as possible by fixed-interest loans. A compro-mise is the convertible bond which affords the lender an equity interest by giving him the right to convert his loan at a future date into shares at a stipulated price. Preference shares are now rarely issued.

(iii) *Government grants and loans*
Grants and loans are available for venture capital (see below), to firms setting up in Assisted Areas (see p. 256), and also to farmers on a percentage basis for expenditure on certain improvements, e.g. woodland and hedge-planting, slurry disposal.

(iv) *Retained profits*
Not all profits are distributed to shareholders. In addition to providing for depreciation and for a contingency fund, profits will be regarded by a successful company as its major source of capital for future expansion.

(c) Financing the expansion of a company
Because the shares of a *private company* are illiquid in that they cannot be offered for sale by public advertisement, a difficult stage in its growth may be reached when its capital is in the region of £250,000, for it is still too small to make a public issue. The gap can be bridged in five main ways.

First, as part of the government's desire to encourage growth of the economy through the development of small businesses, new sources of 'venture capital' have arisen. Banks and other institutions have been more willing to provide medium-term loans especially as, under the Loan Guarantee Scheme, the Department of Industry guarantees 80 per cent of loans up to £75,000. Second, a stockbroker may effect a 'private placing' of shares or debentures with a life insurance company or an investment trust, who are usually in a position to ignore the disadvantages of holding securities of private companies. Third, help might be obtained from the new issue market, where both issuing houses and merchant bankers assist firms to raise capital even providing some themselves. Fourth, there are a number of specialised finance corporations, such as 3i (formerly known as Investors in Industry). Fifth, the commercial banks have now entered the medium-term loan field, particularly through their merchant bank branches.

When a large amount of capital is required, the first step is usually to form a *public company*. But it is the second step which is really important – getting its shares 'quoted' on the Stock Exchange (see pp. 186–90). This entails an exhaustive examination of the company's affairs which have to be advertised very fully in at least two leading London newspapers.

The capital required can be raised by a 'placing', an 'offer for sale' or a 'public issue by prospectus'. The first is the usual method when under £15 million is required, for the costs of underwriting and administration are less. An issuing house, licensed dealer or investment company agrees to sell blocks of the shares direct to institutions and persons who it knows are likely to be interested in them.

For larger amounts an offer for sale is a likely method. The shares are sold *en bloc* to an issuing house, which then offers them for sale by advertisement similar to a public issue.

For more than £50 million, a public issue by prospectus can be employed. Here the company's object is to obtain in a single day the capital it requires. Hence it must advertise well and price its shares a little on the cheap side. The advertisement is in the form of a prospectus which sets out the business, history

(a) Advantages

(b) Disadvantages

(c) Capital source

SOLE TRADER

1. Personal control
2. Personal rewards for energy and initiative

1. Limited capital
2. Unlimited liability
3. Lack of continuity

Personal savings
Loans from Enterprise Allowance Scheme, friends and bank

PARTNERSHIP
2-20 (normally)

1. Additional capital
2. Additional specialised skill

1, 2, 3 above   4. Each partner responsible for actions of others

Partners' savings

PRIVATE COMPANY Ltd.
(one that is not a public company)

1. Limited liability
2. Continuity

1. Corporation tax paid
2. Shares cannot be offered by public subscription
3. Capital still small

Profits ploughed back
'Private placing'
Business Expansion Scheme
Equity Capital for Industry, etc.

PUBLIC COMPANY PLC
2 + shareholders
Share capital of at least £50 000

1. & 2 above
3. Easy to raise capital
4. Easy to expand

1. Accounts have to be published
2. Shareholding and management divorced

Selective marketing: 'placing'
'Offer for sale'
Public issue
'Rights' issue

Figure 5.1   The forms of enterprise and the raising of capital

and prospectus of the company together with its financial standing and the security offered.

In practice, the sale is usually conducted through an *issuing house*, which advises on the terms of the issue. It will also arrange to have the issue underwritten: that is, it will find a number of institutions, such as merchant bankers, which, in return for a small commission, will take whatever part of the issue is left unsold. However, such underwriters do not have to rely entirely on permanent investors to buy the securities on the day of issue, for speculators, known as 'stags', are usually operating, and they buy the shares hoping to resell them quickly at a small profit. Furthermore, where a company is raising additional capital, existing shareholders are now usually given the right to purchase new shares through a *rights issue* in proportion to shares already held and usually at a favourable price.

5.5 The division of labour

(a) Advantages of specialisation

In organising production, the firm will have to consider the advantages of specialisation, the fundamental principle of modern production. Here we examine it with particular reference to labour – although, as we shall see, it is equally applicable to machines, the distribution of goods, localities and even countries.

Where workers are organised so that each specialises on a particular task, increased production results. This is because:

(i) *Each man is employed in the job in which his superiority is most marked*
Suppose that, in one day, Smith can plane the parts for 20 tables *or* cut the joints for 10, whereas Brown can either plane 10 tables *or* cut the joints for 20. If each does both jobs and spends one day one them, their combined production will be 15 tables planed *and* 15 table-joints cut. But Smith is better at planing, while Brown is better at cutting joints. If they specialise on what they can do best, their combined production will be 20 tables planed *and* 20 table joints cut – an increase in output of a third.

Even if, initially, workers were equally proficient at the different jobs, it might still pay to specialise, for the following reasons.

(ii) *Practice makes perfect, and so particular skills are developed through repetition of the same job*

(iii) *Economy in tools allows specialised machinery to be used*
This is illustrated in Figure 5.2, where in (b) division of labour has been introduced. Not only are specialised tools in constant use but their output is much greater. Thus division of labour sets free talented men for research – and allows their inventions to be used profitably.

(iv) *Time is saved through not having to switch from one operation to another, e.g. in obtaining and replacing different tools*
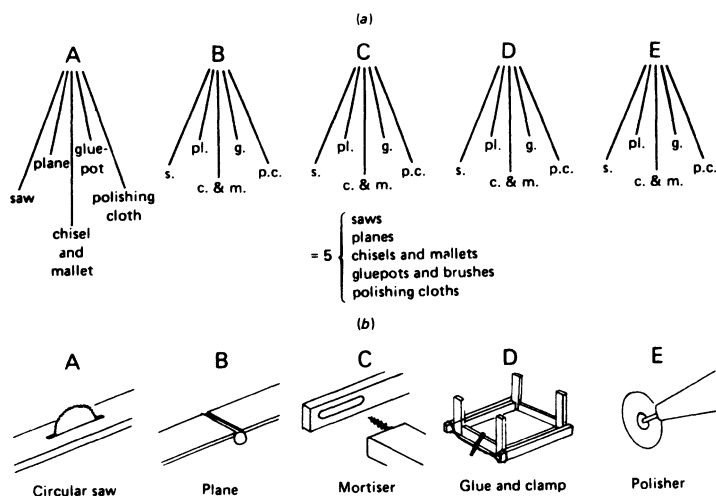
Figure 5.2   Economy in tools through specialisation

(v) *Less time is taken in learning a particular job*

(vi) *The employer can estimate his costs of production and output more accurately*

(b) Disadvantages of the division of labour
While the division of labour leads to lower costs of production, it may have dis-
advantages both for the worker and for society. The worker may find his job
monotonous, and with some occupations such as paint-spraying there is a risk of
occupational disease. Moreover, the skilled specialist may face redundancy if
demand falls, while a strike by a few key workers can lead to widespread unem-
ployment. Finally, standardised products tend to replace individual craft work.

(c) Limitations on the division of labour
Naturally the scope for the division of labour varies from one industry to another.
Countries like Switzerland which have too few workers to permit much special-
isation concentrate on manufacturing a narrow range of products. Again, in
industries such as agriculture and building where the same operations are not tak-
ing place each day, many 'Jacks of all trades' are required. Moreover, an
exchange system using money is essential: we must first unite in exchange before
we can divide in production. Finally, the division of labour has to be related to
current demand for the product. It is no use specialising in making something
which nobody wants; conversely, minute division of labour is only possible when
there is a large demand. The complex organisation of car production, for
instance, rests on a mass demand for a standardised product made up from a mul-
titude of small parts.

5.6   The advantages of large-scale production

As a firm's output increases, costs per unit may fall as a result of the advantages of large-scale production. These are often referred to as 'internal economies' to distinguish them from 'external economies', which arise indirectly from the growth of the *industry* (see p. 71).

(a) Internal economies
Internal economies are of five main kinds:

(i) *Technical economies*
In making a good, as distinct from distributing it, increased output permits more division of labour, greater specialisation of machines, the economy of large machines (e.g. a double-decker bus can carry twice as many passengers as a single-decker, though neither the initial cost nor running costs are doubled) and the linking of processes (e.g. in steel-making, where re-heating is avoided).

Generally technical economies fix the size of the unit actually producing, e.g. a supermarket, rather than the size of the firm, which may consist of many units, e.g. Tesco. Where technical economies are great, the size of the typical unit will tend to be large – as, for example, in the production of cars, sheet steel, gas and electricity. Where, however, increased output merely means duplicating and re-duplicating machines, the tendency will be for the unit to remain small. For instance, in farming at least one combine harvester is necessary for about 600 acres. Thus farms tend to remain small, for as yet there are no greater technical economies to be derived from large machines. Where few technical economies can be gained and yet the firm is large, consisting – as with chain stores – of many operating units, it is usually because other types of economy are possible, as follows.

(ii) *Managerial economies*
When output increases, division of labour can be applied to management. For example, in a shop owned and run by one man, the owner, although having the ability to order supplies, keep accounts and sell the goods, has yet to do such trivial jobs as sweeping the floor, weighing articles and packing parcels – tasks within the capability of a boy who has just left school. His sales, however, may not warrant employing a boy. The large business overcomes this difficulty: a brilliant organiser can devote all his time to organising, the routine jobs being left to lower-paid workers.

The function of management can itself be divided, e.g. into production, sales, transport and personnel. These departments may be further subdivided – sales, for instance, being split into sections for advertising, exports and customers' welfare.

(iii) *Commercial economies*
If a bulk order can be placed for materials and components, the supplier will usually quote a lower price per unit, since this enables him also to gain the advantages of large-scale production.

Economies can also be achieved in selling the product. If the sales staff are not being worked to capacity, the additional output can be sold at little extra cost. Similarly, administrative costs are spread. Indeed, the large firm often manufactures many products, so that one acts as an advertisement for the others. Thus Hoover vacuum-cleaners advertise their washing-machines, dishwashers and steam-irons. In addition a large firm may be able to sell its by-products, something which might be unprofitable for a small firm.

Finally, when the business is sufficiently large, the division of labour can be introduced on the commercial side, with expert buyers and sellers being employed.

Such commercial economies represent real gains to the community, reducing prices through better use of resources. On the other hand, where a large firm uses its muscle to *force* suppliers into granting it favourable prices, it will simply result in higher prices to other buyers.

(iv) *Financial economies*

In raising finance for expansion the large firm is in a favourable position. It can, for instance, offer better security to bankers – and, because it is well-known, raise money at lower cost, since investors have confidence in and prefer shares which can be readily sold on the stock exchange.

(v) *Risk-bearing economies*

Here we can distinguish three sorts of risk. First, there are risks which can be insured against, enabling large and small firms alike to spread risks.

Second, certain businesses usually bear some risk themselves, saving some of the profits made by the insurance company. Here the large firm has a definite advantage. London Transport, for instance, can cover its own risks, while a large bank can call in funds from other branches when there is a run on the reserves in a particular locality.

The third kind of risk is one that cannot be reduced to a mathematical probability and thus cannot be insured against – risk arising from changes in demand for the product or in the supply of raw materials; this is usually referred to as risk arising from 'uncertainty'. To meet fluctuations in demand the large firm can diversify output (like British American Tobacco) or develop export markets. On the supply side, materials may be obtained from different sources to guard against crop failures, strikes, etc.

(b) External economies

While the firm can plan its internal economies, it can only *hope* to benefit from external economies which arise as the *industry* grows.

First, the concentration of similar firms in an area may produce mutual benefits: a skilled labour force; cooperation in common services such as marketing and research; better roads and social amenities; technical schools catering for the local industry; product reputation; ancillary firms supplying specialised machinery, collecting by-products, etc. The firm must take into account such economies when deciding where production shall take place, for the lower costs may outweigh any diseconomies which arise through traffic congestion, smoke, etc.
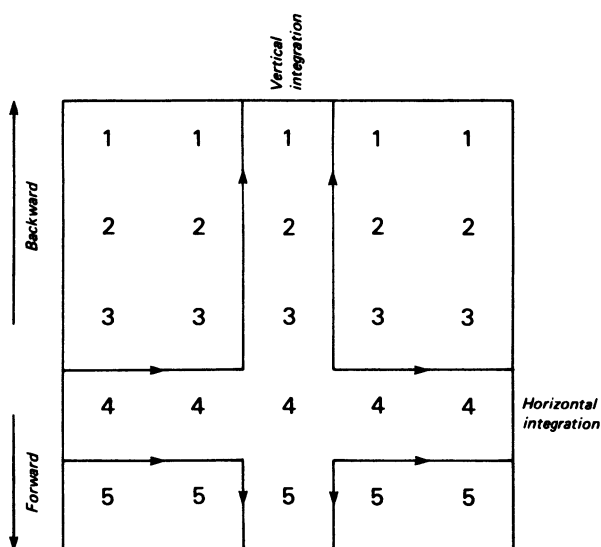
Figure 5.3 Horizontal and vertical integration

Second, external economies can take the form of common information services provided either by associations of firms or even by the government.

Finally, as an industry grows in size, specialist firms may be established to provide components for all producers thereby extending economies of scale.

5.7   The size of firms

(a) Horizontal, vertical and lateral combination
The advantages of large-scale production provide firms with a strong impetus to combine.

Horizontal integration occurs where firms producing the same type of product combine. Thus Nestlé took over Rowntree, and Ford acquired Jaguar.

Vertical integration is the amalgamation of firms engaged in the different stages of production of a good. Thus Britoil, an oil exploration company was taken over by BP. Vertical integration may be 'backward' towards the raw material, or 'forward' towards the finished product.

Both the above can improve efficiency, thereby lowering costs per unit, and increasing profits. Thus horizontal integration can allow greater specialisation, commercial economies and a saving on administrative overheads. Vertical integration facilitates linked processes and reduces risk by increasing direct control over the supply and quality of raw materials and components. Moreover, all parts can be manufactured to an integrated design, and there is direct control over the distribution of the final product (see below).

Lateral integration occurs where a firm increases the range of its products. Concentration on one product may make a firm vulnerable to a change in fashion, a switch in government policy or a recession. Thus the firm diversifies, often by

taking over other firms producing completely different products. For instance, P & O is engaged in shipping, cross-channel ferries, road transport and construction through its subsidiary companies.

Apart from increased efficiency and security of profits, integration may enhance a firm's prestige. One other aim, however, must not be overlooked – monopoly power. This is discussed in Chapter 6.

Integration may result from internal development or combining with existing firms either by merger or a takeover – when a company buys all the shares of a smaller firm and absorbs it completely – or by the formation of a holding company in which the parent company obtains enough shares to give it effective control, though the smaller company preserves its identity and enjoys considerable independence of action. Many large companies, e.g. Unilever, GEC and Great Universal Stores, hold such controlling interests in subsidiary companies.

(b) The predominance of the small firm

In spite of the advantages enjoyed by the large firm, we must not conclude that every firm has to be large to be competitive. Indeed the small firm still predominates in all forms of production. In agriculture two-thirds of all holdings are less than fifty hectares in size, while in retailing nearly three-quarters of all firms consist of only one shop. Even more remarkably, the same is true of manufacturing, where one would have thought that technical economies of scale would be all-important. Table 5.1, which shows the size of the establishment – the factory or workshop – in manufacturing, reveals two important features: (i) the small establishment is typical of manufacturing in the UK, over nine-tenths employing less than 100 people; (ii) these small units employ only one-third of the labour force.

Any explanation of this predominance of the small firm has to deal with two salient facts: (i) small firms are especially important in certain industries, such as agriculture, retailing, building, and personal and professional services; (ii) variations in the size of firms exist even within the same industry. Both result from the nature of the conditions of demand and supply.

Table 5.1   Size of manufacturing establishments in the UK, 1992

Employees	Number of firms	Percentage of total firms	Number of employees (000s)	Percentage of total employed
1–9	99 767	65.3	325	7.1
10–99	43 946	28.8	1 322	28.6
100–999	8 728	5.7	2 200	47.6
Over 1000	383	0.2	771	16.7
Total	152 824	100	4 618	100

Source: *Annual Abstract of Statistics.*

(i) *Demand*

Large-scale production may be only *technically* efficient; it is not *economically* efficient unless a large and regular demand justifies it.

The market may be small because demand is local (e.g. for personal services and the goods sold by the village store), or limited to a few articles of one pattern (e.g. for prestige luxury goods and highly specialised and individually designed machine-tools) or because transport costs are high (e.g. for bricks and perishable market-garden produce), or because product differentiation divides it artificially (see p. 94).

Where demand fluctuates (e.g. in construction), the overhead cost of idle specialised equipment is heavy – but the smaller the firm, the less the burden.

(ii) *Supply*

Even if demand is large, factors on the supply side may make for small firms. While in certain industries, e.g. retailing and building, it is possible to start with little capital, or be supported by franchising (e.g. McDonald's) or by joining a wholesale chain (e.g. Spar), the difficulty of obtaining further funds and the taxation of profits are obstacles to expansion. Furthermore, government monopoly policy may prevent mergers (see p. 110). Alternatively, where vertical disintegration is possible, firms need not expand internally but simply employ specialist firms for advertising, research, supplying components and selling by-products. Important, too, is the fact that many small owners do not have the drive to expand or the ability to manage a large concern. Or, as in farming and retailing, they will work long hours (that is, accept a lower rate of profit) simply to be their own bosses.

Above all, as the size of the firm increases, management difficulties occur. If management is vested in heads of department, problems of co-ordination arise and rivalries develop. This means that one person must be in overall command – yet people with such capabilities are in very limited supply. In certain industries these difficulties may soon occur. Rapid decisions are required where demand changes quickly, e.g. in the fashion trades, or supply conditions alter, e.g. through the weather in agriculture. Or care may have to be given to the personal requirements of customers, e.g. in retailing and services. This may necessitate the close supervision of management, and thus the firm has to be small.

5.8 The location of production

In deciding where to produce, a firm has to weigh the advantages of a particular locality against the rent or land costs it will have to pay there compared with elsewhere.

(a) Location advantages

The advantages of production in a particular locality can be classified as: (i) natural, (ii) acquired and (iii) government-sponsored.
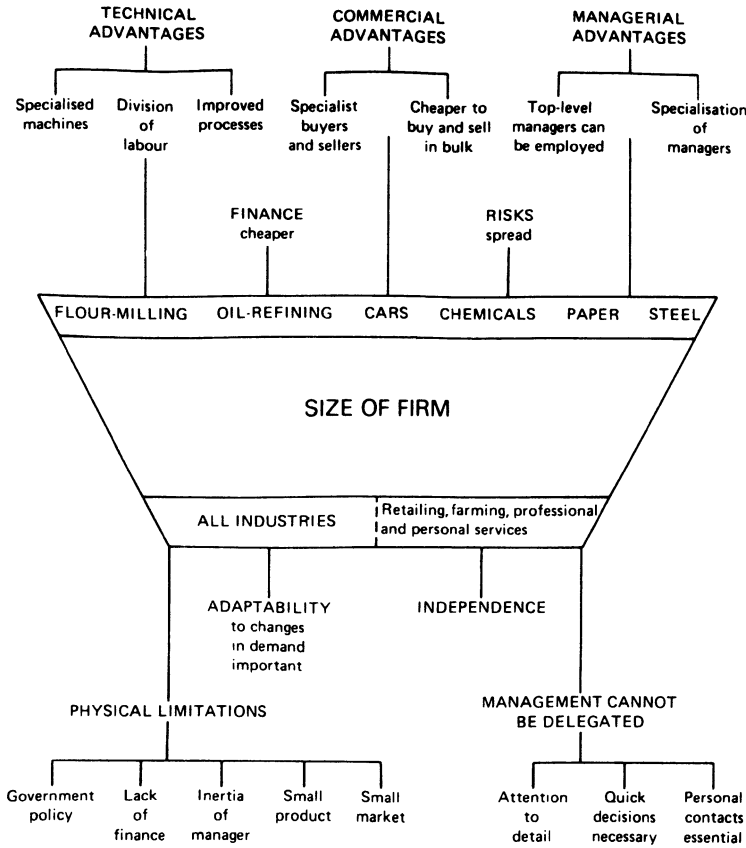
Figure 5.4 Factors influencing the size of the firm

(i) *Natural advantages*

Costs are incurred both in assembling raw materials and in distributing the finished product. With some goods the weight of the raw materials is far greater than that of the finished product. This is particularly true where coal is used for heat and power, e.g. in iron and steel production (Figure 5.5).

Here transport costs are saved by producing where raw materials are found, e.g. on coal- and iron-ore fields, or where they are easily accessible, e.g. near a port.

On the other hand, in some industries the cost of transporting the finished product are greater than those of assembling the raw materials, e.g. with ice-cream, mineral waters, furniture and metal cans. With these it is cheaper for a firm to produce near the market for its goods. Thus Walls has ice-cream factories close to most large concentrations of population.

What is really important as regards transport costs is their ratio to the value of the product. Thus sand and gravel are excavated locally, whereas special types of brick are transported long distances.

Coal

1¾ tons

Limestone

¼ ton

Iron ore

2 tons

Blast furnace

1 ton
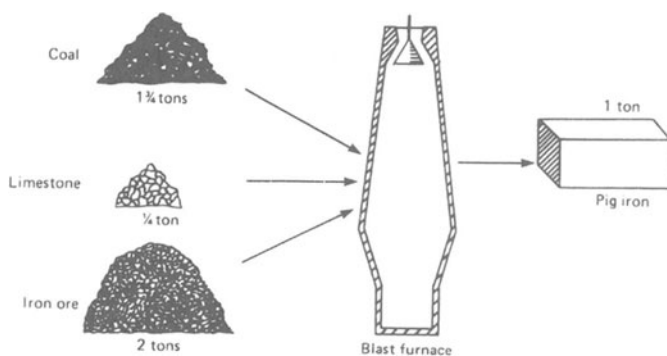
Pig iron

Figure 5.5   The production of pig-iron – a 'weight-losing' industry

Generally speaking, transport improvements and new developments (e.g. electrical power, lighter materials) have helped firms to move away from their sources of raw materials. The tendency now is, therefore, for firms to concentrate, not on the coalfields, but on the outskirts of areas of high population which provide both a supply of labour and a market for the finished good.

A river, estuary or coastal location may be essential when huge quantities of water are required by an industry (e.g. chemicals, atomic power), and this may also be important for waste disposal.

Besides accessibility to raw materials and nearness of markets, suitability of climate is a further natural advantage which may have to be considered when locating production. Indeed, in agriculture, it is usually decisive, provided soil conditions are not adverse.

Under 'natural advantages' we can also include an adequate supply of the type of labour required. Thus high technology industries have been attracted to the south-east of England by the skilled labour available, while the abundance of cheap labour has been important for the development of mass-production in Taiwan and Hong Kong.

(ii) *Acquired advantages*
Improved methods of production, the development of transport, inventions and new sources of power may alter the relative importance of natural advantages and so change an industry's location. Thus, as high-grade iron-ore fields have become exhausted and improved techniques have reduced coal consumption, it is now cheaper to transport the coal than the iron ore to produce pig-iron, and so the industry has shifted to the ports importing iron ore and to the low-grade iron-ore fields of the east Midlands. Similarly, improved transport may upset the relative pulls. Finally, new inventions, such as humidifiers and water-softeners (cotton and wool), can make an industry less dependent upon a particular locality.

Yet we must not over-stress the importance of the above changes. Even when natural factors have disappeared, an industry often remains in the same region because of the 'man-made' advantages it has acquired, e.g. steel, cotton. Such advantages were mentioned earlier when we studied external economies of

concentration. A skilled labour force, communications, marketing and commercial organisations, nearby ancillary industries (to achieve economies of scale or to market by-products), training schools and a widespread reputation for the products of the region all help to lower the costs of production, thereby making the locality attractive to new firms.

(ii) *Government-sponsored advantages*

Unemployment in such highly localised industries as coal, cotton and ship-building, and environmental problems (traffic congestion, pollution, housing stress) in regions attracting new and expanding industries, have led the government to offer firms financial inducements to set up plants in Assisted Areas (see Chapter 21).

(b) The level of rents in different areas

Location advantages have to be weighed against the cost of land (or, where it is hired, rent). This cost varies from one locality to another and is determined by the market. Since other firms, possibly from other industries, may be looking for the same site advantages, competition will fix the price of land at the highest which the keenest firm is prepared to pay. This will be the firm which puts the greatest value on the land's advantages compared with those of land elsewhere. Thus, early in its history it seemed that the cotton industry might settle on Clydeside, for this had all the natural advantages of south-east Lancashire. But it also had advantages for producing iron and steel and for building ships, and in these her superiority was most marked. Thus shipbuilding firms were prepared to pay extra for this advantage. For cotton manufacturers this extra cost of a site on the Clyde exceeded any disadvantage of being in Lancashire. Thus shipbuilding firms settled along the Clyde, and cotton firms in Lancashire.

   In the final analysis, therefore, it is not the absolute advantages of a district which decide where a firm locates, but the advantages relative to those of other districts. Thus an industry whose outlay on unskilled labour forms a high proportion of its production costs would, other things being equal, be able to bid more for land in an area of cheap labour than one whose spending on such labour was minimal. And in town centres we see the same principle at work – shops oust other businesses, and houses are converted into offices.

(c) Other influences on location

A firm will normally choose the site where the advantages are greatest compared with its cost. But even for a comparatively new industry, where natural advantages are important, we cannot assume that they will be decisive. Thus it is largely historical accident which accounts for the presence of the Rover plant at Cowley on the outskirts of Oxford, for the old school of William Morris came up for sale just as the production of cars at his original cycle works was being expanded.

   Moreover, electricity has now practically eliminated dependence on a coalfield site. Yet firms may still go to the original areas because of the advantages acquired over time. Others may choose to be nearer their markets. Some 'foot-

loose' firms have even located in certain districts, particularly south-east England, largely because the managing directors (or their wives) have preferred living there!

The various factors influencing location are summarised in Figure 5.6.
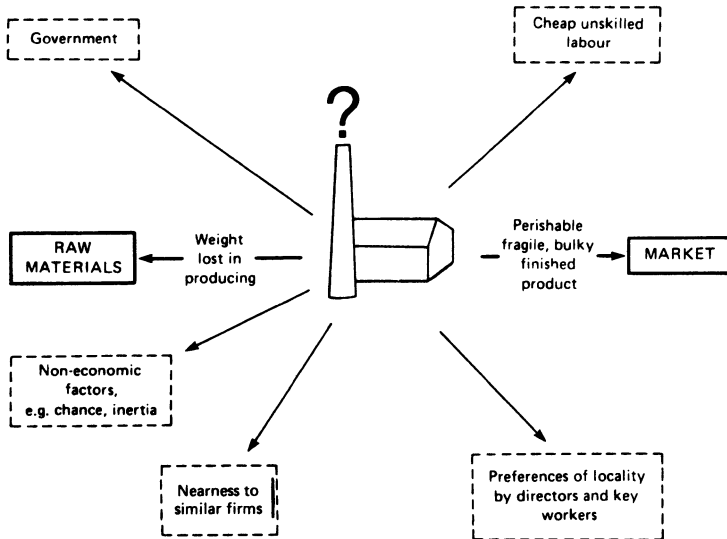
Figure 5.6   Factors influencing the siting of a business

5.9   The distribution of goods to the consumer

(a) The scope of production

A manufacturer has to decide how to get his finished goods to the consumer. He may undertake the task himself. But if he does so, he must employ salesmen, run delivery transport, carry stocks, advertise his product, organise exports, advise customers,establish servicing centres and give credit. Experts for these highly specialised functions can only be employed if output is large enough. Moreover, the manufacturer's main ability lies in organising production rather than its distribution.

Thus the principle of the division of labour is usually applied. Just as the manufacturer buys raw materials and components from other producers, so specialist firms get his goods to the consumer – there is *forward* vertical dis-integration. We will simplify our account of this selling process by grouping such firms into 'wholesalers' and 'retailers'. Figure 5.7 shows their place in the various stages of the production of chocolate.

(b) The wholesaler

The wholesaler buys goods in bulk from producers and sells them in small quantities to retailers. In doing so he helps production in a number of ways.
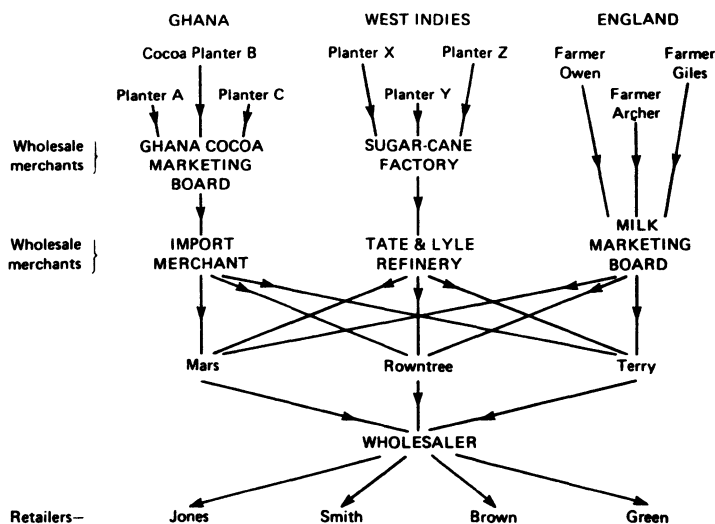
Figure 5.7   The role of the wholesaler and retailers in the production of chocolate

(i) *He economises in distribution*

Because shops usually stock a variety of goods, they can order supplies only in small quantities. Thus it is not economical for each producer to sell directly to them.

Figure 5.8 shows that, when the chocolate firms deliver in bulk to a wholesaler, the number of contacts and journeys is reduced from sixteen (a) to eight (b).

Particularly in agriculture, where the goods are perishable, the farmer finds it easier if, instead of trying to contact retailers himself, he delivers his produce to a wholesaler or commission salesman, for example at New Covent Garden, and leaves the actual selling to him.

Similarly, in the construction industry, where there are numerous small builders, it is easier for manufacturers to deliver through builders' merchants.

(ii) *He keeps stocks*

Consumers like the convenience of being able to obtain a good at a shop just when they require it. This means that stocks have to be held. Often, however, neither the producer nor the retailer has the necessary storage facilities or the extra capital required, and so it is left to the wholesaler.

In other ways, too, the costs of storage are removed from the producer or retailer. While loss through fire, flood or rats can be insured against, no insurance can be taken out to cover a fall in demand. Thus a wholesaler or dealer, who holds stocks of a good which is liable to go out of fashion, relieves manufacturers and retailers of that risk.

The holding of stocks is, in itself, a valuable economic function in that they help to even out fluctuations in price resulting from temporary fluctuations in
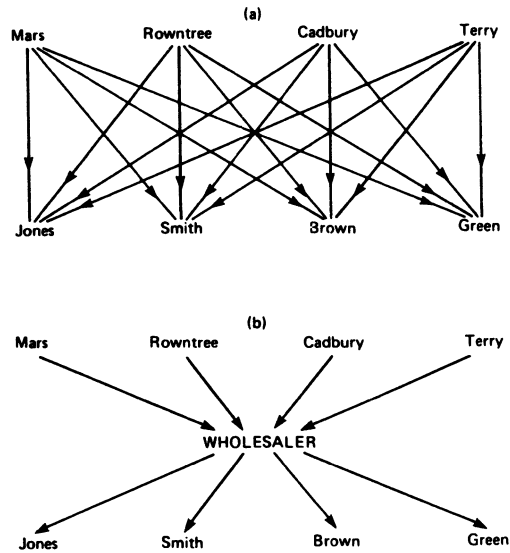
Figure 5.8   Economising in distribution through the wholesaler

demand or supply. Thus merchants replenish brick stocks during the winter, and run them down during the rest of the year.

(iii) *He arranges imports from abroad*
Manufacturers could rarely be bothered to ship small parcels to individual retailers abroad or to undertake the currency and documentary arrangements. They prefer to deal with a wholesaler, an import merchant with established trade connections.

(iv) *He carries out certain specialised functions*
Not only does the wholesaler advertise goods but, in order to make selling easier, he may process goods – pasteurising milk, blending tea, refining sugar, and grading commodities such as wheat, cotton and wool tops.

(v) *He is a channel for information and advice*
Suggestions which customers make to the retailer are passed on to the wholesaler, who, if he sees that they reflect the general view, conveys them to the manufacturer. Thus the latter can improve his product and anticipate fashion changes.

(vi) *He assists in the day-to-day maintenance of the good*
With many products, particularly vehicles and machinery, the wholesaler relieves the manufacturer of the task of providing an efficient maintenance, repair and spare-part service.

(c) The retailer

The retailer performs the last stage of the productive process, for it is he who puts the goods in the hands of the actual consumer. His work is to have the right goods in the right place at the right time.

(i) *He stocks small quantities of a variety of goods*

What the 'right good' is depends on the customer, for different people have different tastes. It is therefore necessary to stock a variety of goods so that customers can choose and take delivery there and then. Thus the retail shop is basically a showroom, particularly where goods are bought infrequently.

   The size of the stocks carried will depend on many factors: the popularity of the product, the possibility of obtaining further supplies quickly, the perishability of the good or the likelihood of its going out of fashion, the season, the possibility of future price changes and, above all, the cost (chiefly bank overdraft interest) of carrying stocks.

(ii) *He takes the goods to where it is most convenient for the customer*

Taking the goods to the customer usually means that the retailer sets up his shop within easy reach, e.g. in a town centre. However, with goods in everyday use, such as groceries, small shops are often dotted around residential districts. Where customers are very dispersed, the retailer may even be a 'travelling shop'.

   While customers take most goods away with them, the retailer arranges delivery if transport is essential, e.g. for furniture, or if the customer likes the extra convenience of delivery, e.g. of the morning milk.

(iii) *He performs special services for customers*

In the course of his main business the retailer performs many services to build up customer goodwill. Where the good is not in stock he will order it, and in other matters where contact with the manufacturer is necessary he often acts for the customer, e.g. by returning the goods for repair.

   With many goods, too, such as fishing tackle, photographic equipment, musical instruments and machinery, he can usually provide specialised advice.

   Finally, goods may be sent on approval, or finance arranged through hire purchase, special credit accounts, etc.

(iv) *He advises the wholesaler and manufacturer of customers' preferences*

(d) Types of retail outlet

Retailing might be widely defined as including all shops, mail-order firms, garages, launderettes and betting shops, and indeed any business selling products or services to the consumer. However, it is usual to confine the meaning of 'retailing' to shops and mail-order outlets, as follows.

(i) *Independents*

These are mainly small shops with no branches, and they account for nearly two-fifths of total shop sales. Yet, in spite of their advantages of individual attention to customers, handy locations for quick shopping trips, and the willingness of own-

ers to accept a lower return in order to be their own bosses, these independents have steadily lost ground to the larger stores.

A major bid to avert the decline has come through the voluntary chains, such as Spar, Mace and Wavy Line, of which about a third of the independents are members. While retaining their independence, members buy in bulk from wholesalers and use common advertising and display techniques.

A more recent development enabling an individual trader to set up in business is 'franchising', that is buying a concession from a major firm, e.g. Kentucky Fried Chicken, McDonalds, 'K' shoes, United Dairies, to sell the brand-name product. The concessionaire advertises and advises on presentation.

(ii) *Multiples*
These can be defined arbitrarily as organisations of ten or more shops. Some, such as Mothercare and Dixon's, sell a particular type of good. Others, such as Woolworth, Boots and Marks & Spencer, have a fairly extensive range of products. Together with the supermarkets they comprise about 48 per cent of the retail market.

Their chief advantages are economies of bulk-buying and centralised control, the elimination of the wholesaler, quick recognition through standardised shop-fronts, and a reputation established through brand names.

(iii) *Supermarkets*
These may be defined as self-service shops with a minimum selling area of 200 sq. metres, but the trend in recent years, especially among the multiple grocery companies, has been towards superstores (of at least 2000 sq. metres) and hypermarkets (see below). While organisationally they would count as multiples, their share of the food trade warrants separate attention. In 1991 they accounted for half the grocery trade and over a third of retail food sales.

The field is led by the five major retail grocery chains: Tesco, Sainsbury, Associated Dairies (ASDA), Argyll Stores and Gateway. Their strength lies in economies of scale, low labour costs, a clear and attractive display of merchandise, bulk buying and selling under their own label (e.g. Sainsbury cornflakes, Tesco coffee). As a result they have highly competitive prices and have gained ground rapidly.

Indeed, many of these self-service organisations have extended their activities beyond groceries to self-service of goods showing higher profit margins, e.g. clothing, hardware and do-it-yourself and garden supplies.

(iv) *Hypermarkets*
Urban congestion, inadequate parking space and rising rents have made high street sites increasingly expensive. The answer to these problems has been the very large 'out-of-town' shopping centre or 'hypermarket', catering mainly for the car-borne weekly shopper, and these outlets are now expanding rapidly.

(v) *Department stores*
Competition from multiples has forced department stores to alter somewhat their traditional pattern of having separate departments under the control of a

responsible buyer, often described as 'many stores under one roof', in favour of bulk-buying by central office, more self-service, and extended credit facilities. As a result they have retained some 4 per cent of the retail market, but the modern development is towards covered shopping malls.

The main groups are House of Fraser, Debenhams, the John Lewis Partnership and Great Universal Stores.

(vi) *Co-operatives* (see p. 62).

(vi) *Mail order*
Mail-order business, which accounts for nearly 3 per cent of the retail market, is particularly susceptible to higher postal charges.

The major companies, Great Universal Stores, Littlewoods Mail Order, Freemans and Empire Stores, sell by agency and illustrated catalogues, purchases usually being arranged through weekly interest-free payments. Over one-half of all sales are accounted for by women's clothing and household goods.

(e) Factors affecting the type of retail outlet
Over the last thirty years, the pattern of retailing has moved away from the small, independent shop towards the larger organisation, notably the multiples, supermarket chains and mail-order firms. This trend reflects a greater emphasis on competition through lower price rather than by better service.

The larger firms are in a strong position to cut prices. Not only do they obtain the advantages of large-scale production (particularly those of selling a whole range of goods and of buying in bulk), but they can use their bargaining strength to secure further price discounts from manufacturers. Indeed, the largest may force the manufacturer to supply goods under the retailer's 'own-brand' label at a price below that at which other retailers can buy the manufacturer's national brand. Moreover, since large retailers cater for a whole range of shopping, e.g. food, they can attract customers into stores by 'loss-leaders'.

Economic factors influencing this trend have been:

   (i) *Increased income,* which has led to a swing towards the more expensive processed foods and consumer durable goods and facilitated less frequent shopping trips.
  (ii) *An increase in car ownership*, which has enabled people to move from the city centre to the outer suburbs. Shops have followed, not only to be near their customers, but also to obtain larger sites with parking facilities, lower rents and less congestion. The car has also made customers more mobile, enabling them to travel to good shopping centres where they can purchase all their requirements at a single stop.
 (iii) *An increase in the number of married women going to work*, which has promoted the demand for convenience foods and labour-saving devices. It has also led to the reduction of the number of shopping expeditions, a trend helped by the wider ownership of refrigerators and freezers.

These factors are likely to remain important in the future. It seems probable, therefore, that new supermarkets will take the form of discount stores or hyper-

markets selling a wider range of products having a higher profit margin than groceries. Moreover, the more favourable response to recent planning applications is enabling new stores to be developed outside towns, while cash-and-carry warehouses are now available to consumers who can buy in quantity.

Such changes are likely to be at the expense of the medium-sized business, for the smaller local retailers offer 'convenience' services.

(f) The future of the middleman

Wholesalers and dealers who come between the manufacturer and the retailer or consumer are often referred to as 'middlemen'. They are frequently criticised on the grounds that they take too large a share of the selling price. It is argued that, if the manufacturer sold direct to consumers, prices could be reduced.

But, as we have seen, wholesalers relieve producers of essential functions, allowing them to obtain the advantages of specialisation in marketing products. Such forward vertical dis-integration is the cheapest way of getting goods to the consumer.

However, this does not mean that all criticism of middlemen is unjustified. Sometimes their profit margins are too high. This may occur through continuing with antiquated methods or by a single middleman's playing off one small producer, such as a farmer, against another (hence the formation of producers' co-operatives).

In recent years a tendency for the wholesaler to be eliminated has been due to: (i) the growth of large shops, which can order in bulk; (ii) the development of road transport, which reduces the necessity of holding large stocks; (iii) the desire of manufacturers to retain some control over retailing outlets in order to ensure that their products are pushed or a high standard of service, freshness, etc. is maintained; and (iv) the practice of branding many products, which eliminates many specialised functions. In other cases, however, the elimination of the wholesaler has been confined to sales of high-value goods, such as furniture and television sets; to circumstances where the producer and retailer are close together, as with the market gardener who supplies the local shop; and to cases where the manufacturer does his own retailing.

To some extent the wholesaler has responded to this challenge by developing in two main directions: (i) by establishing the cash-and-carry warehouse, sometimes called 'the retailers' supermarket'; and (ii) by becoming the organiser of a voluntary chain of retailers, who are supplied, and to some extent controlled, by him, e.g. Spar.

Selling direct to consumers by the manufacturer occurs chiefly where: (i) the manufacturer wishes to push his product (e.g. beer or footwear) or to ensure a standard of advice and service (e.g. sewing-machines); (ii) the personal-service element is important (e.g. made-to-measure clothing); (iii) the manufacturer is a small-scale producer-retailer, often selling a perishable good (e.g. cakes and pastries) or serving a local area (e.g. with printing); (iv) so wide a range of goods is produced that a whole chain of shops can be fully stocked (e.g. Whitbread's ales, Clark's shoes); or (v) the good is highly technical or made to individual specifications (e.g. machinery).

⬡ 6 Supply: (ii) costs and profitability

6.1 Combining resources

(a) Classification of factors of production

In order to examine the problems connected with employing resources it is helpful to classify them according to particular characteristics.

Land refers to the resources provided by nature, e.g. space, sunshine, rain and minerals, which are fixed in supply.

Labour refers to the effort, physical and mental, made by human beings in production. It is this 'human' element which distinguishes it from other factors, for it gives rise to problems regarding psychological attitudes and unemployment.

Capital as a factor of production consists of producer goods and stocks of consumer goods not yet in the hands of the consumer. While consumer goods directly satisfy consumers' wants, e.g. loaves, bicycles, TV sets, producer goods are only wanted for making consumer goods, e.g. buildings, machines, raw materials. Capital is treated as a separate factor of production in order to emphasise: (i) the increased production which results from using it, (ii) the sacrifice of present enjoyment which is necessary to obtain it (see Chapter 11), and (iii) the fluctuations of economic activity which may result from changes in its rate of accumulation (see Chapter 18).

Enterprise is the acceptance of the risks of uncertainty in production – risks which, as we saw earlier, cannot be insured against. They arise because the firm spends in advance on raw materials, labour and machines, and the extent to which such costs are covered depends on the demand for the product when it is sold. Tastes may have changed or a rival may be marketing the good at a lower price than anticipated. The reward for uncertainty-bearing is profit – unless it is negative: loss. Whoever accepts such a risk is a true entrepreneur – the farmer working on his own account, the person who buys ordinary shares in a company, or the citizen of a state, who ultimately has to bear any losses made by a nationalised industry.

(b) The problem of combining resources

The problems peculiar to the different types of factors of production are considered in Chapters 9–11. Here we are concerned with the more general problem of how much of each a firm will hire. In other words, how will the firm allocate its spending in order to obtain the greatest possible output from a given outlay? For example, the same amount of concrete can be mixed by having many men with just a shovel apiece or by having only one man using a concrete-mixer. Can we discover any general principle governing the firm's decision? We can begin by

seeing what happens to output when one factor is held fixed while the amount of another factor is increased.

(c) The law of diminishing (or non-proportional) returns

Assume: (i) production is by two factors only, land and labour; (ii) all units of the variable factor, labour, are equally efficient; (iii) there is no change in techniques or organisation.

Table 6.1 shows how the output of potatoes varies as more labourers work on a fixed amount of land. Until 3 men are employed, the marginal product of

Table 6.1   Variations in output of potatoes resulting from a change in labour employed

Number of men employed on the fixed unit of land	*Total output*	*Output (50 kg bags)*	
		Average product	*Marginal product*
1	2	2.0	2
2	16	8.0	14
3	54	18.0	38
4	80	20.0	26
5	95	19.0	15
6	108	18.0	13
7	120	17.1	12
8	130	16.2	10
9	138	15.3	8
10	142	14.2	4
11	142	12.9	0

Notes:
  (a) *Total output* is the total output (bags) from all factors employed.
  (b) *Average product* refers to the average output per man. It therefore equals

$$\frac{\text{total output}}{\text{number of men employed}}$$

  (c) *Marginal product* refers to the marginal output (bags) to labour, and equals the addition to total output which is obtained by increasing the labour force by one man. That is, marginal output equals total output of $(n + 1)$ men minus total output of n men.
  (d) There is a fundamental relationship between average product and marginal product. Marginal product equals average product when the latter is at a maximum (fig 6.1). This relationship is bound to occur. So long as the marginal product is greater than average product, the return to an additional labourer will raise the average product of all labourers employed. On the other hand, as soon as the marginal product falls below average product, the additional labourer will lower the average product. Hence when average product is neither rising nor falling, that is, at its maximum, it is because marginal product equals average product.
       This relationshp can be made clearer by a simple example. Suppose Atherton has played 20 innings and that his batting average is 60 runs. Now if in his next innnings he scores more than 60, say 102, his average will increase – to 62. If, on the other hand, he scores less than 60, say 18, his average will remain unchanged at 60.

labour is increasing – the third labourer, for instance, adding 38 bags. Here there are really too few labourers for the given amount of land. Thereafter the marginal product falls, the fourth labourer adding only 26 bags, and so on; total output is still increasing, but at a diminishing rate. The maximum return per labourer occurs when there are 4 labourers to the plot. If we increase the number of labourers to eight, the maximum return per labourer can only be maintained by doubling the amount of land. When 11 labourers are employed they start to get in one another's way, and from then on total output is declining absolutely.

Again it must be emphasised that units of the variable factor are homogeneous. The marginal product of labour does not fall because less efficient labourers have to be employed. Diminishing returns are the result of more labourers being employed on a fixed amount of land.

Nor does the law formulate any *economic* theory; it merely states physical relationships. While the physical productivity of an extra labourer is important to a farmer in deciding how many men to employ, it will not *determine* his decision. He must also know the relative costs of factors; that is, he requires economic data as well as technical facts.

(d) The practical applications of the law of diminishing returns

The law is significant both in our everyday life and in the theoretical analysis of the economist.

First, it helps to explain the low standard of living in many parts of the world, particularly the Far East. Increasing population is cultivating a fixed amount of
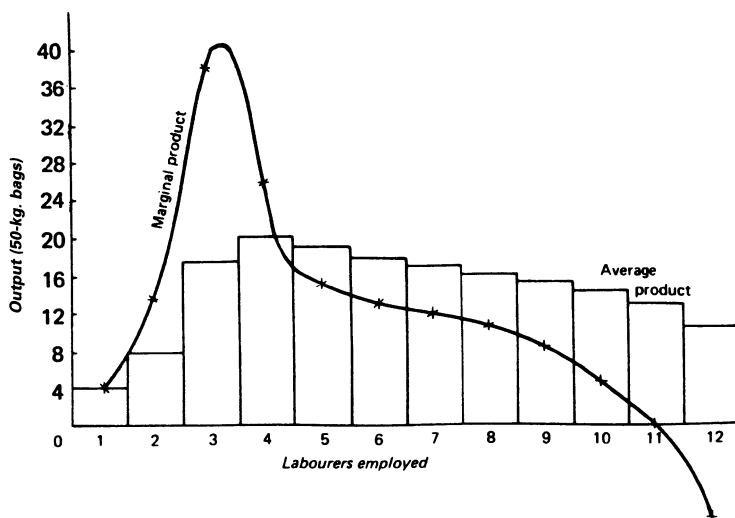

Figure 6.1   The relationship between the number of labourers employed
average product and marginal product

land. Marginal product, and thus average product, are falling; so, therefore, is the average standard of living.

Second, it shows how a firm can adjust the marginal physical products of factors by altering the proportion in which they are combined. Thus few labourers to the plot gave a high return per labourer; after 4 labourers, the average product began to fall. So the law is often referred to as 'the law of varying proportions'. The firm will choose that combination of factors which yields the maximum output from a given outlay, as follows.

(e) The optimum combination of variable factors

So far we have assumed that there are just two factors, land and labour, and that land is fixed. But suppose that there is another variable factor, say capital. Now the farmer will have to decide how he will combine labour with capital.

How much of each he employs will depend upon its productivity relative to its price, since he will alter the combination until, for the last pound spent on both labour and capital, he obtains the same amount of product. Suppose, for instance, the last pound's worth of labour is yielding more potatoes than the last pound spent on capital. It will obviously pay the farmer to transfer this pound from capital to buying more labour, for this will increase his total physical yield.

But labour and capital are obtained in different units, their units being different in price. Thus we cannot directly compare the productivity of one man with that of one unit of capital, say a mechanical hoe; we must allow for their respective prices. If the cost of one man is only one-third of the cost of a mechanical hoe, then the marginal product of a man need only be one-third of the hoe's to give the same yield for a given expenditure. Thus the farmer will be in equilibrium in combining factors when:

$$\frac{\text{Marginal product of labour}}{\text{Price of labour}} = \frac{\text{Marginal product of capital}}{\text{Price of capital}}$$

A corollary of this is that, like the housewife in purchasing her goods, the firm will tend to buy more of a factor as its price falls, and less as it rises. Suppose the wage-rate rises but the marginal product of labour remains unchanged. The fundamental relationship stated above has now been destroyed. To restore the position it is necessary to raise the marginal product of labour and to lower that of capital by combining less labour with more capital: in short, a rise in wages without a corresponding increase in the productivity of labour will, other things being equal, tend to bring about the replacement of labour by machines (see p. 142)

6.2 The costs of production

(a) Opportunity costs and profit

Suppose a man sets himself up as a shopkeeper selling sweets. He invests £4,000 of his savings in the business, and in the first year his receipts are £40,000 and his

outgoings £22,000. The accountant would say that his profits over the year were £18,000. The economist, however, would disagree.

The reason for this is that the economist is not so much concerned with money costs as with 'opportunity cost' – what a factor could earn in its best alternative line of production. This concept of cost has a bearing on (i) the economist's concept of 'profit', and (ii) how long production should continue when total costs are not covered.

(b) 'Implicit costs'

The £22,000 money outgoings of the shopkeeper above can be regarded as 'explicit costs'. But when we look at costs as alternatives forgone we see immediately that the shopkeeper has certain 'implicit costs' – the rewards his own capital and labour could earn elsewhere. If, for instance, his capital could be invested at 6%, there is an implicit cost of £240 a year. Similarly with his own labour. His next most profitable line, we will assume, is to be a shop manager earning £11,760 a year. Thus a total of £12,000 in implicit costs in addition to the explicit costs should be deducted from his receipts.

(c) Normal and super-normal profit

But we have not finished yet. The shopkeeper knows that, even in running a sweet shop, some risk arises through uncertainty – a risk which he avoids if he merely works for somebody else. The shopkeeper must therefore anticipate at least a certain minimum profit, say £2,000 a year, before he will start his own business. If he does not make this minimum profit, he feels he might as well go into some other line of business or become a paid shop manager. Thus another type of cost (which we call 'normal profit') has to be allowed for – the minimum return which keeps a firm in a particular industry after all other factors have been paid their opportunity cost. Normal profit is a cost because, if it is not met, the supply of entrepreneurship to that particular line of business dries up.

We have, therefore, the following costs: explicit costs, implicit costs and normal profit. Anything left over after all these costs have been met is 'super-normal' profit. In terms of our example, we have:

		£	£
Total revenue			40,000
Total costs:	explicit	22,000	
	implicit	12,000	
	normal profit	2,000	
			36,000
Super-normal profit			4,000

(d) Fixed costs and variable costs

For the purposes of our analysis, we shall classify costs as either *fixed* or *variable*.

Fixed costs are those costs which do not vary in direct proportion to the firm's output. They are the costs of indivisible factors, e.g. buildings, machinery and

vehicles. Even if there is no output fixed costs must be incurred, but for a time, as output expands, they remain the same.

Variable costs, on the other hand, are those costs which vary directly with output. They are the costs of the variable factors, e.g. operative labour, raw materials, fuel for running the machines, wear and tear on equipment. Where there is no output, variable costs are nil; as output increases so variable costs increase.

In practice it is difficult to draw an absolute line between fixed and variable costs: the difference really depends on the length of time involved. When current output is not profitable, the entrepreneur will have to contract production. At first overtime will cease; later, workers will be paid off. In time, more factors, e.g. salesmen, become variable, and if receipts still do not justify expenditure on them they too can be dismissed. A factor becomes variable when a decision has to be taken on whether it shall be replaced, for then its alternative uses have to be considered. Eventually machines need renewing; even they have become a variable cost. A decision may now be necessary on whether the business should continue.

The distinction between fixed and variable factors and costs is useful in two ways. First, in economic analysis it provides a means of distinguishing between differences in the conditions of supply which result from changes in the time period. The *short period* is defined as a period when there is at least one fixed factor. While, therefore, supply can be adjusted by labour working overtime and more raw materials being used, the time is too short for altering fixed plant and organisation. Thus the firm cannot achieve its best possible combination for a given output. In the *long period* all factors are variable; they can therefore be combined in the best possible way. Supply can now respond fully to a change in demand.

Second, as we shall see later, the distinction is fundamental when the firm is considering whether or not to continue producing. In the long period all costs of production, fixed and variable, must be covered if production is to continue. But in the short period fixed costs cannot be avoided by ceasing to produce; they have already been paid for, simply because it was necessary to have some 'lumpy' factors even before production could start. Only variable costs can be saved; and so, provided these are covered by receipts, the firm will continue to produce. Anything that it makes above such costs will help to recoup its fixed costs.

(e) Changes in costs as output expands

In discussing the law of diminishing returns we referred to quantities of factors and yields in physical terms. But in deciding how to maximise profit, the firm will be concerned with those quantities translated into money terms. It can then see directly the relationship between costs and receipts at different outputs and is thus able to decide what output will give the maximum profit (see Table 6.2). Our first task, therefore, is to consider how costs are likely to change as output increases. We shall assume perfect competition in buying factors of production – the demand of each firm is so small in relation to total supply that any change in demand will not directly affect the price of those factors.

In the short period there are, by definition, bound to be fixed factors. And when considering the law of diminishing returns we found that when a variable factor was added to a fixed factor the marginal product might increase for a time but would eventually diminish. How will this affect costs as output expands?

Let us assume that two factors are being used, one of them fixed. If each unit of the variable factor costs the same, but the output from additional units is increasing, the firm is obtaining an increasing amount of output for any given addition to expenditure. In other words the cost of each additional unit of output is falling as output expands. On the other hand, if the marginal product of the variable factor if diminishing, the cost of an additional unit of output is rising. This cost of an additional unit of output is known as *marginal cost* (MC).

The above conclusions are represented diagrammatically in Figure 6.2, where average product = total product of x units of the variable factor/x and average variable cost = total variable costs of n units of output/n.
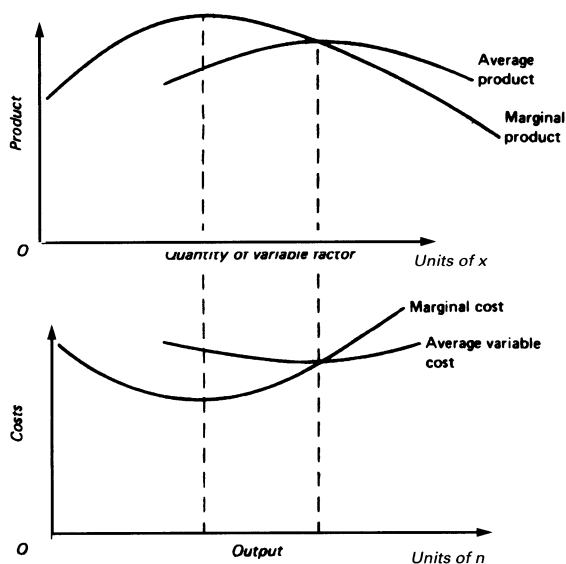

Figure 6.2   The relationship between returns and costs

(f) Costs schedules
Table 6.2 illustrates this relationship between output and costs. The figures, which have been kept as simple as possible, are for an imaginary firm, Roller-mowers, manufacturers of lawnmowers. Fixed costs (FC) amount to £10,000, and, as variable factors are added, output expands. At first there is an increasing marginal product; as a result MC is falling. This has its effect on average total

Table 6.2   Costs of Rollermowers (in £)

Output per week (units)	Fixed cost (FC)	Total variable cost (TVC)	Total cost (TC)	Marginal cost (MC)	Average fixed cost (AFC)	Average variable cost (AVC)	Average total cost (ATC)
0	10000	–	–	–	–	–	–
				200			
10	10000	2000	12000		1000	200	1200
				140			
20	10000	3400	13400		500	170	670
				100			
30	10000	4400	14400		333	146.7	480
				100			
40	10000	5400	15400		250	135	385
				135			
50	10000	6750	16750		200	135	335
				185			
60	10000	8600	18600		167	143.3	310
				240			
70	10000	11000	21000		142.9	157.1	300
				300			
80	10000	14000	24000		125	175	300
				390			
90	10000	17900	27900		111.1	198.9	310
				510			
100	10000	23000	33000		100	230	330
				660			
110	10000	29600	39600		91.1	269.1	360
				840			
120	10000	38000	48000		85	316.7	400

Notes:

   (1)   TC of n units = FC + VC of n units.

   (2)   MC is the extra cost involved in producing an additional unit of output. That is, MC of the nth unit = TC of n units – TC of $n-1$ units. Here output is shown in units of 10, so that this difference in total costs has to be divided by 10.

   (3)   AFC of n units = $\dfrac{FC}{n}$.

   (4)   AVC of n units = $\dfrac{TVC \text{ of } n \text{ units}}{n}$.

   (5)   ATC = $\dfrac{TC \text{ of } n \text{ units}}{n}$.
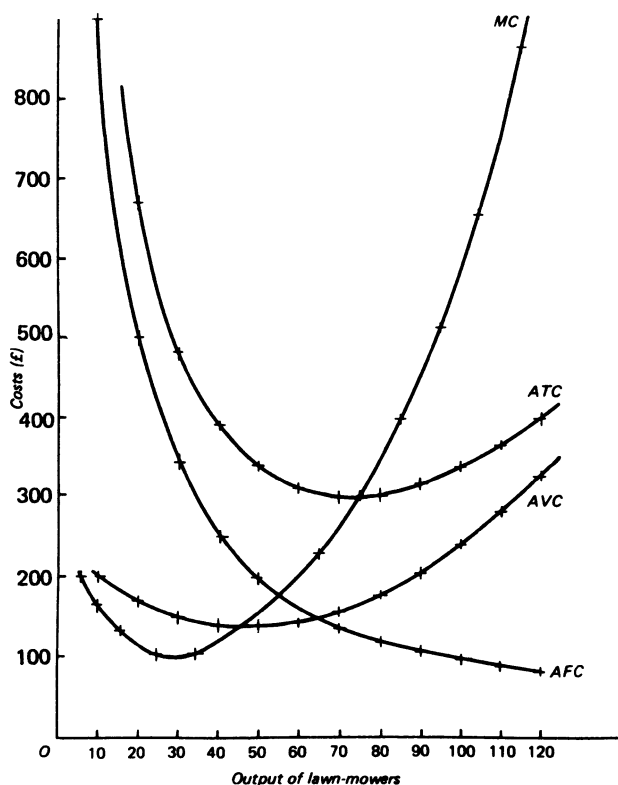
Figure 6.3   Cost curves

cost (ATC) until approximately 75 units are being produced. From then onwards, as the fixed factors are being worked more intensively, diminishing returns cause the ATC curve to rise. These figures can be plotted on a graph (Figure 6.3).
   The following relationships between the curves should be noted:
  (i)   AFC and AVC added vertically give ATC
  (ii) The MC curve cuts both the AVC and ATC curves when they are at a minimum, the same reason applying as in Table 6.1, note (d).

6.3   How much to produce: output of the firm under perfect competition

In order to ascertain whether a firm is maximising profits, we have to know (i) the price at which it can sell different outputs and the price at which it can buy different quantities of factors, and (ii) whether it is free to enter another industry where it can make higher profits. Both depend on the extent to which competition prevails.
   First we shall assume that the conditions of 'perfect competition' – the highest form of competition – apply. Later we show how relaxing these conditions leads to imperfect competition, forms of which prevail in real life.

(a) The conditions necessary for perfect competition
For perfect competition to exist the following conditions must hold:

(i) *A large number of relatively small sellers and buyers*
If there are a large number of sellers relative to demand in the market, any one seller will know that, because he supplies so small a quantity of the total output, he can increase or decrease his output without having any significant effect on the market price of the product. In short he is a 'price-taker', and can sell any quantity at this price.

This is illustrated in Figure 6.4, where (a) shows market price *OP* determined by the demand for and supply of the goods of the industry as a whole. But the industry supply, we will assume, comes from a thousand producers, each of about the same size. Each producer therefore sells such a small proportion of the total market supply that he can double his output from *ON* to *OM* or halve it from *OM* to *ON* without affecting the price – Fig 6.4b.

In other words, in perfect competition a seller is faced with an infinitely elastic demand curve for his product. If, in our example, he charges a higher price than *OP*, nobody will buy from him; if he charges less than *OP*, he will not be max- imising his revenue, for he could have sold all his output at the higher price, *OP*.

In contrast, the producer in Figure 6.5b sells such a large proportion of the market supply that a change in his output affects the price he receives for his product. When he supplies *OM*, the price is *OP*. If he increases his supply to *OM*₁, the price falls to *OP*₁. Similarly, if he decreases his supply to *OM*₂, the price rises to *OP*₂. Alternatively such a producer can decide on the price he charges, leaving it to the market to determine how much is sold at that price. But he can- not fix both price and quantity at the same time.

(ii) *Homogeneous product*
Buyers must regard the product of one producer as being a perfect substitute for that of another, and purchase solely on the basis of price, switching to a competitor if one producer raises his price.
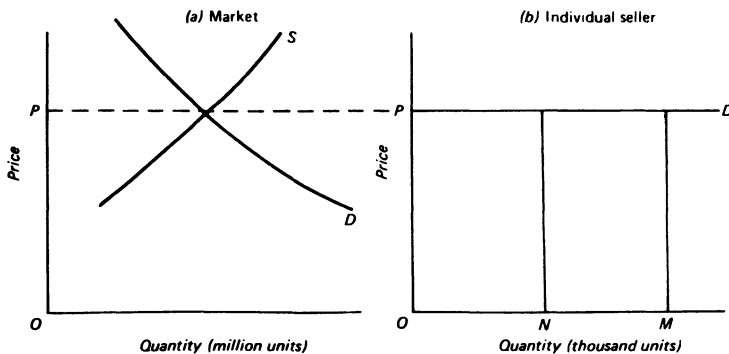

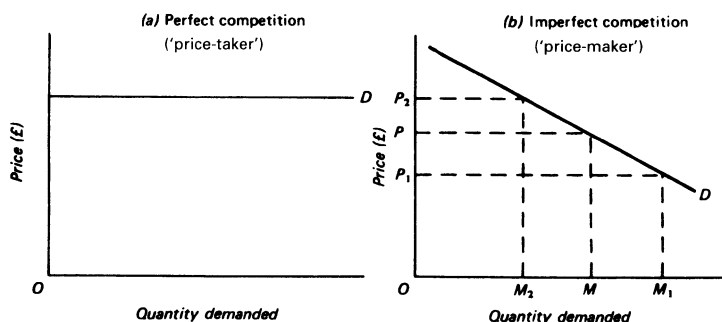**Figure 6.4 The firm's demand curve under perfect competition**

Figure 6.5 The firm's demand curve under perfect and imperfect competition

Such identity of product does not exist where there is a real or imaginary dif-ference (e.g. a special wrapping or brand name) or where reasons other than price (e.g. goodwill) influence buyers. Here an individual producer can raise his price without necessarily losing all his customers. In short, product differentiation leads to some downward slope in the demand curve.

(iii) *A perfect market, especially perfect knowledge of market conditions*
There are two aspects of perfect knowledge:

(a) sellers and buyers must know the prices being asked in other parts of the market, both product and factor, so that they can act accordingly;
(b) in order to make free entry effective, a would-be producer must also know what profits are being made by existing producers.

The above conditions give a perfectly competitive market. For a situation of perfect competition to exist we must also have:

(iv) *Free entry of new firms into the market*

(v) *Perfect mobility of the factors of production in the long period*
A change in the demand for a product must, in the long period, result in the transfer of factors of production from one line of production to another.
    In practice these conditions never apply simultaneously, and perfect competi-tion must be regarded primarily as an analytical device which enables us to arrive at some fundamental conclusions.

(b) Maximising profit
Since the objective of the firm, we have assumed, is to maximise its profits, it will seek to produce that output where the difference between total revenue and total costs is greatest. The firm, therefore, will be concerned with two broad questions: (i) How much will it obtain by selling various quantities of its product? (ii) How much will it cost to produce these different quantities?

At first sight it may seem that maximum profit will occur at the minimum average cost output. But this is unlikely to be so. The real question which the entrepreneur will be continually asking is: 'If I produce another unit, will it cost me less or more than the extra revenue I shall receive from the sale of it?' That is, he concentrates his attention at the margin: if an extra unit of output is to be profitable, *marginal revenue* (the addition to total revenue received from it) must at least equal *marginal cost* (the cost of producing it).

Under perfect competition the firm will obtain the market price for its goods, whatever its output. In other words, marginal revenue (MR) equals price, with the price line horizontal (Figure 6.5 and 6.6). On the other hand, although under perfect competition the firm can buy increasing quantities of its factors at a given price, the MC curve eventually rises because of diminishing returns.

(c) The equilibrium output of Rollermowers

Let us return to our imaginary firm. Assume that the market price of mowers is £450. We can impose this MR curve on the cost curve diagram (Figure 6.6).

Now at any output where MR (price) is above MC, Rollermowers can increase profits by expanding output. Or, if MC is above MR, contracting output will
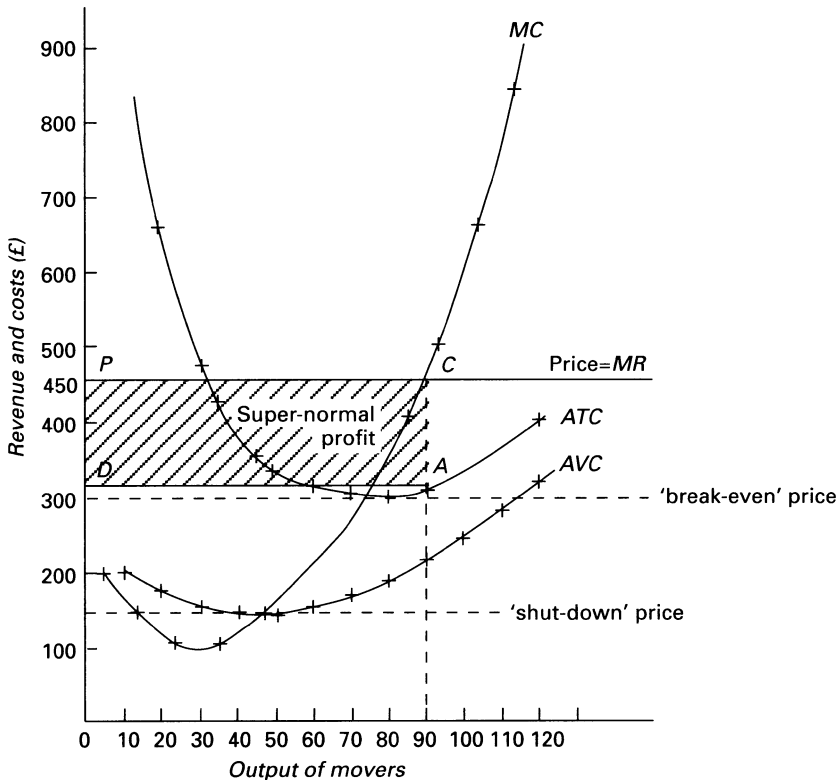

Figure 6.6   The equilibrium output of the firm under perfect competition

increase profits. The equilibrium output, therefore, is where MR (price) equals MC, that is, 90 lawnmowers. Here average total cost is £310. Thus super-normal profit equals total receipts (£40,500) − total costs (£27,900) = £12,600 = shaded area *PDAC*.

But, for any output, current revenue must cover current costs. And while 'current revenue' is simply the number of goods currently produced times their price, 'current costs' depend as we have seen, on whether we are dealing with the short or the long period.

(d) The short-period 'shut-down' price
A firm will only *start* to produce if it expects that total revenue will be sufficient to cover (i) the cost of fixed factors, (ii) the cost of variable factors, e.g. labour, raw materials, and (iii) normal profit.

We will imagine that the firm does think it can make a go of it. It buys highly specific machinery (fixed costs) which, we will assume for the sake of simplicity, has no value to any other firm, together with labour and raw materials (variable costs), and starts producing.

But as time goes by it finds that its original expectations are not being fulfilled. The price at which it can sell its good is lower than estimated. Although the cost of variable factors is being covered, the firm sees that, unless price rises the margin between the two is too small to provide sufficient cash to replace machines when they wear out. In other words the business as a whole will prove unprofitable.

But what will our firm save by stopping production forthwith? Obviously its variable costs, for these vary directly with output. But what of its machines, which, since they have no alternative use, have no resale price? These are fixed factors which have already been paid for, and ceasing to use them now cannot recoup past expenditure. Their opportunity cost is zero.

Consequently our firm takes a philosophic view of the situation. It has some perfectly good machines which, if used, will add nothing to costs. So, provided the cost of the variable factors is being covered, it goes on producing. Anything earned above such cost will help to recoup the cost of the fixed factors.

How can we tell if variable costs are being covered? Simply by looking at the AVC curve. If we take Rollermowers as an example, a price of £135 for a mower would just enable it to produce in the short period. Here MC would equal MR and, with an output of 45 units, TVC would just be covered. Any price lower than this, however, would mean that, for any output where MC = MR, total receipts (price times output) would be less than TVC (AVC times output). Rollermowers could not make a 'go' of it even in the short period; and so we can call £135 the 'shut-down' price.

(e) The firm's short-period supply curve
A firm's MC curve is its short-period supply curve. At any price below £135 per mower, Rollermowers will stop production, because TVC are not covered. At higher prices, however, it will produce an output where price equals MC, as follows:

Price (£)	Output (units)
135	45
185	55
240	65
300	75
390	85, and so on.

This schedule is graphed in Figure 6.7.

6.4 The industry's supply curve

(a) The short-period supply curve
In the short period, no new firms can enter the industry because, by definition, they cannot obtain fixed factors. The supply curve of the industry, therefore, is obtained simply by adding the output of all existing firms at each given price.

Suppose, for the sake of simplicity, that the industry consists of four firms, the other three being less efficient than Rollermowers. Their outputs (starting from minimum AVC) are given under A, B and C in Table 6.3.

Table 6.3   Short-period supply schedule

Price (£)	Output (units)				
	Firm A	Firm B	Firm C	Rollermowers	Total
135	–	–	–	45	45
185	–	–	45	55	100
240	–	45	55	65	165
300	50	55	65	75	245
390	55	65	75	85	280
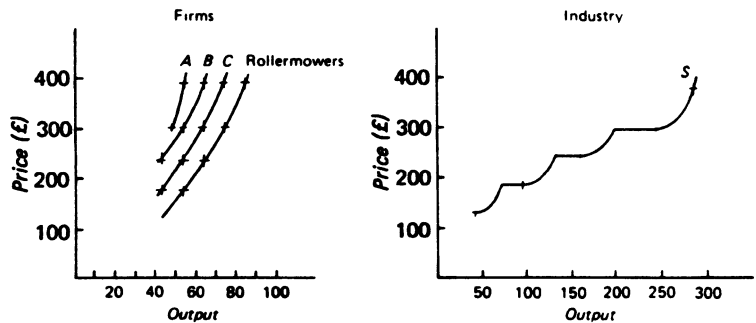
Figure 6.7 The short-period supply curve of the industry

This is shown graphically in Figure 6.7. The MC curves of the four firms are summed horizontally to obtain the short-period supply curve of the industry. This rises from left to right, showing that more is supplied the higher the price.

It will be observed that the supply curve derived above is not smooth, but stepped. This is because we have taken only four firms. If there had been many firms, each differing only slightly in efficiency, we would have had a smoother curve.

(b) Supply in the long period

In the long period a firm will still produce where MR = MC, but total costs must be covered.

As regards the industry, however, any super-normal profits being made will attract new firms, for these can now obtain plant. Moreover, competition will force all firms, both old and new, towards the most efficient size. As a result supply increases and the price of the product falls until super-normal profits are eliminated (Figure 6.8).

The above argument would produce a horizontal supply curve. In practice, however, it has to be modified. First, while there may be some external economies as the industry expands, there is a major diseconomy – higher rewards have to be paid to attract factors from other industries. Second, entrepreneurs are unlikely to be equally efficient in looking ahead when making their decisions, and some firms will always be doing better than others. The situation, therefore,
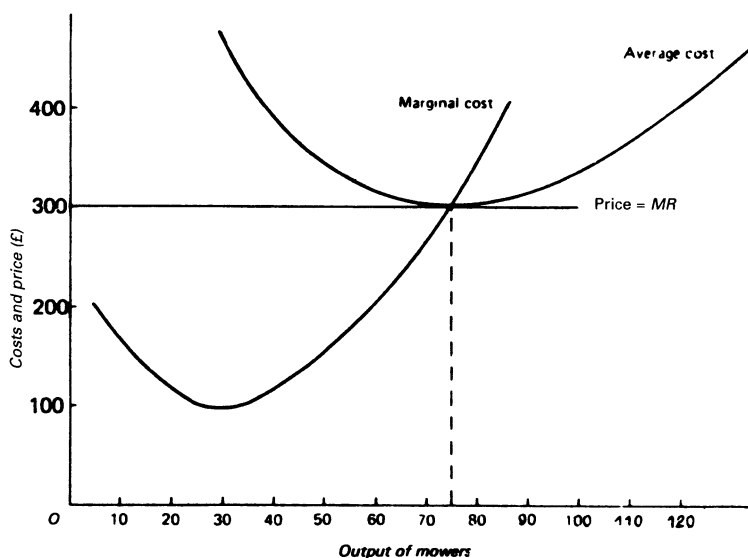
Figure 6.8   The effect of competition on the super-normal profits and output of Rollermowers

is that even in the long period there is likely to be an upward-sloping supply curve. The extent to which this happens is indicated by elasticity of supply.

6.5   Elasticity of supply

(a) Definition

Consider Figure 6.9. For a rise in price from OP to OP_1, supply extends from OM to OM_1 with S_1 and to OM_2 with S_2. At price OP, therefore, S_2 is said to be more elastic than S_1.

More precisely, the elasticity of supply of a good at any price or at any output is the proportional change in the amount supplied in response to a small change in price divided by the proportional change in price. In the supply schedule on p. 37, for instance, when the price of eggs rises from 10p to 12p supply expands from 32,000 to 40,000. Elasticity of supply is therefore equal to:

$$\frac{\dfrac{8}{32}}{\dfrac{2}{10}} = \frac{5}{4} = 1.25$$

As with elasticity of demand, we say that supply at a given price is elastic if elasticity is greater than 1, and that it is inelastic if elasticity is less than 1.
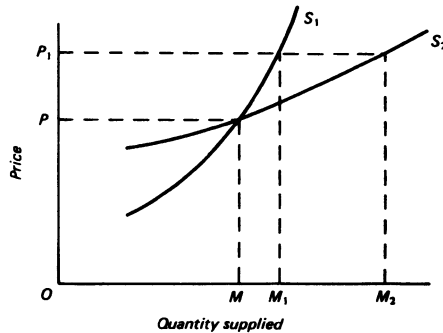
Figure 6.9   Elasticity of supply

There are two significant limiting cases.

(i) *Elasticity of supply equal to infinity*
The main uses of this concept are: (1) where there is perfect competition in buying factors of production; and (2) where production takes place at constant cost. In both cases the supply curve is horizontal – Figure 6.10a.
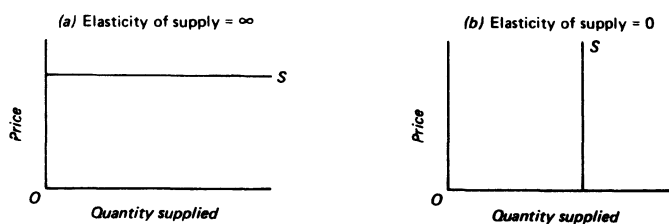
Figure 6.10   Extremes of elasticity of supply

(ii) *Supply absolutely inelastic*
Here a good is fixed in supply whatever the price offered – Figure 6.10b. This applies to rare first editions and old masters, and by definition to fixed factors in the short period.

(b) Factors determining elasticity of supply
Elasticity of supply is determined by: (i) the period of time under consideration; (ii) the relationship between the individual firms' minimum-supply points; and (iii) the cost of attracting factors from alternative uses. We shall consider each in turn.

(i) *Time*
We distinguish three main periods:

(1) MOMENTARY EQUILIBRIUM
Here the supply is fixed, and elasticity of supply = 0. An example is Christmas trees on Christmas Eve. With many goods, some increase in supply can take place by drawing on stocks, utilising any spare capacity, or switching factors of production from one product to another (where a firm makes two or more different products).

(2) SHORT-PERIOD EQUILIBRIUM
Usually varying supply requires a change in the factors of production employed. But this takes time – and the period differs for each factor. In the short period, as we have seen, it is possible to adjust supply only by altering the variable factors (raw materials, labour, etc.).

(3) LONG-PERIOD EQUILIBRIUM
Other factors – the fixed factors, e.g. land already sown and capital equipment – can be altered in the long period, allowing supply to adjust fully to a change in price. Thus elasticity is greater in the long period. For example, in Figure 6.9, S_1 could well represent the short-period supply curve, and S_2 the long.

(ii) *The relationship between the firms' minimum-supply points*
The supply curve is obtained by aggregating the supply of individual firms. If these firms each offer a supply to the market at more or less the same mini-

mum price, supply will tend to be elastic at that price. Similarly, as price rises, the greater the number of firms coming in, the greater is the elasticity of supply.

(iii) *The cost of attracting factors of production*
In order to expand production additional factors have to be attracted from other industries. For an industry as a whole, this means that higher rewards will have to be paid. What we have to ask, therefore, is how much of a factor will be forthcoming in response to a given price rise. In other words, what is the elasticity of supply of factors of production? And, of greater significance, what influences determine this elasticity?

In answering this question we can first consider what happens when one particular industry, e.g. office-building, wishes to expand. Let us concentrate on one factor: labour. With increased demand for building labourers, wages rise. But they rise not only for the office-building industry but for all other industries employing such labourers – house-building, road-construction, public works, etc. How will it affect these industries?

First, they will try to substitute other factors, e.g. cement-mixers, bull-dozers, etc., for the labour which now costs more. Is such substitution physically possible? If so, is the supply of these alternative factors elastic, or will their prices rise sharply as demand increases? If physical substitution is fairly easy and the supply of alternative factors is elastic, it will mean that a small rise in wages will release much labour for the office-building industry.

Second, higher wages will lead to increased costs in building houses, constructing roads, etc. The supply curve of these products, therefore, moves to the left; and, the higher the proportion of wages to total costs, the further will it move. The extent to which this leads to a reduced production of these alternative goods will depend upon the elasticity of demand for them. If elasticity is high, the small rise in the price of the good will cause a considerable contraction of demand, and labour will be released for office-building. If, on the other hand, demand is inelastic, even a considerable rise in wages will have little effect on the output of houses, etc., and the increase in the supply of labour to office-building will be correspondingly small.

We see, therefore, that the two main influences affecting the elasticity of supply of a factor to a particular industry are (1) the extent to which other factors can be substituted, and (2) the elasticity of demand for the alternative goods it produces.

(c) Practical uses of the concept of elasticity of supply

(i) *The elasticity of supply of a good is a major factor in determining how much its price will alter when there is a change in the conditions of demand*
This can be seen by considering the likely effect on the price of cane sugar, in the short and long period, of an increase in the demand for sugar.

We can assume a fairly inelastic demand curve for sugar. The original price is *OP* (Figure 6.11). Demand then increases from D to D_1. The supply of cane sugar
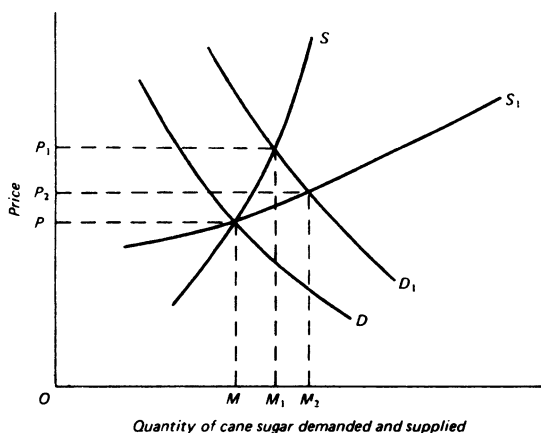
Figure 6.11 Changes in the price of cane sugar over time in response to a change in demand

in the short run is inelastic, for supply can be expanded only by adding labour, fertilisers, etc. Price therefore increases to OP_1. But in the long period more land can be planted with sugar-cane. Supply is now more elastic, and is represented by the curve S_1. The long-run price falls to OP_2.

(ii) *The elasticity of supply is significant with regard to taxation*
First, where the supply of a good is inelastic, the Chancellor of the Exchequer can impose a tax on the producer without its having a great effect on the amount of the good offered for sale. Suppose, for instance, that a man owns a field which is suitable only for sheep-grazing, and that the most any farmer will pay him for the use of this field is £10 a year, which the owner accepts. Now suppose that the government puts a tax of £5 a year on this type of land. This means that the owner will have to pay the tax out of his own pocket, for the farmer will pay no more, and the land cannot be put to any other use. In fact the government could tax almost all the rent away before it would make any difference to the number of sheep being grazed on it. However, if all the rent went on tax, the owner might leave the land standing idle (see p. 152).

Second, the relative elasticities of demand and supply determine the propor-tion of a selective indirect tax borne by the producer as compared with the consumer (see pp. 273–6).

6.6 Monopoly

(a) Imperfect competition
Where any of the conditions of perfect competition are not fulfilled, some form of 'imperfect competition' results and the firm's demand curve is downward-sloping:

(i) A seller may be so large that the quantity he supplies affects the price.
(ii) Products may not be homogeneous, because product differentiation or goodwill allow a producer to raise his price somewhat while still retaining some customers.
(iii) Lack of knowledge, barriers to entry or immobility of factors of production result in imperfect elasticity of demand or supply. Consumers, for instance, may not have complete knowledge of prices ruling elsewhere – as, for example, in retail markets. Thus sellers can raise their prices without losing all their custom.

There are many 'shades' of imperfect competition. At one extreme we have a single producer of a product, e.g. British Oxygen; at the other, the only difference from perfect competition is that firms each produce a slightly different brand, e.g. toothpaste. The first situation we call 'monopoly', the second 'monopolistic competition'. In between we can have just a few sellers of the same or of a slightly different product – 'oligopoly'. In this case, each seller has to take into account the reactions of rivals to his own pricing policy. For instance, if he raises his price, will rival firms follow his lead and do likewise, or will they keep their prices unchanged in order to win over some of his customers? Each different assumption gives its own solution to the equilibrium output of the industry.

The broad market forms are shown in Figure 6.12. We restrict our study to an examination of monopoly.
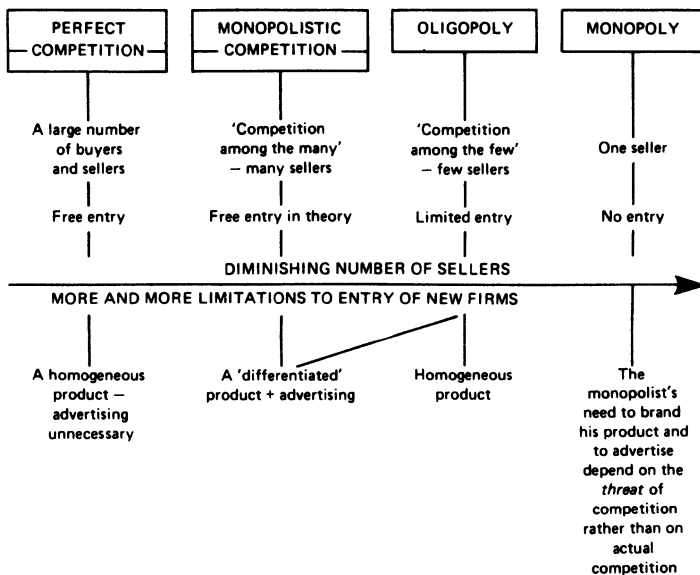

Figure 6.12 Market forms

(b) Sources of monopoly power

While to some extent all goods are substitutes for each other, there may be essential characteristics in a good or group of goods which give rise to gaps in the chain of substitution. If one producer can so exclude competitors that he controls the supply of a good, he can be said to be a 'monopolist'– a single seller.

In real life there is seldom complete monopoly. But one producer may dominate the supply of a good or group of goods. In the UK any firm which accounts for a quarter (in value) of market sales is considered to be a monopoly (see p. 110).

Possible sources of a monopolist's power to exclude competitors are:

(i) *Immobility of the factors of production*

Such immobility means that existing suppliers cannot be challenged by new entrants. It may arise through:

(1) *Legal prohibition of new entrants* – as with certain public utility undertakings, where many firms would create technical difficulties, e.g. natural gas transmission, letter delivery (Post Office).

(2) *Patents, copyrights and trademarks*, where the object is to promote invention and the development of new ideas.

(3) *Government policy of establishing single buying and selling agencies*, e.g. marketing boards.

(4) *Control of the source of supply by one firm,* e.g. mineral springs, diamonds (De Beers), trade unions and professional associations.

(5) *Restrictions on imports* by tariffs, quotas, health controls, etc.

(ii) *Ignorance*

A monopoly may persist largely through the ignorance of possible competitors. They may not know about the super-normal profits being made by the existing firm, or they may be unable to acquire the necessary know-how, e.g. for special technical processes.

(iii) *Indivisibilities*

Whereas the original firm may have been able to build up its size gradually, new firms may find it difficult to raise the large capital required to produce on a scale which is cost-competitive, e.g. with cars, drugs, aircraft.

In some cases, too, the efficient scale of plant may be so large relative to the market that there is only room for one firm. These 'natural' monopolies cover many public utilities, e.g. gas supply, water, electricity generation.

(iv) *A deliberate policy of excluding competitors*

Restriction of competition falls into two main groups. On the one hand we have the sources of monopoly power described so far. These have, as it were, resulted indirectly rather than from any deliberate action by producers. Such 'spontaneous' monopolies must be contrasted with 'deliberate' monopolies – those which are created specifically to restrict supply (e.g. OPEC).

It is essential to distinguish between the two when formulating policy. While the 'spontaneous' monopolies may still abuse their fortunate position in order to make high profits, to a large extent they are inevitable, and usually policy should seek to control rather than destroy them. On the other hand, monopolies solely designed to follow restrictive practices detrimental to the consumer should, where possible, be broken up. In practice, however, it is often difficult to draw a distinct line between the two. While firms may increase production or combine in order to reduce costs through economies of scale, the effect may still be that competitors are forced out.

Deliberate action to exclude competitors takes various forms. Firms producing or selling the same good may combine, or a competitor may be subject to a take-over bid. Monopolies are often formed in the sale of services. Trade unions are primarily combinations of workers formed with the object of obtaining higher wages (see Chapter 10). Certain professions, such as medicine and the law, have their own associations which regulate qualifications for entry, professional conduct, and often fees.

Some practices designed to exclude competitors are highly questionable – vicious temporary price-cutting, collusion in submitting tenders, collective boycotts, intimidation of rivals' customers by threats to cut off the supply of another vital product, etc.

(c) The effect of the downward-sloping demand curve on marginal revenue

Consider Figure 6.13. In (a) the producer is selling under conditions of perfect competition. His marginal revenue is equal to the full price, since all units sell at this. Thus, for the fourth unit, MR is the shaded area *A*.
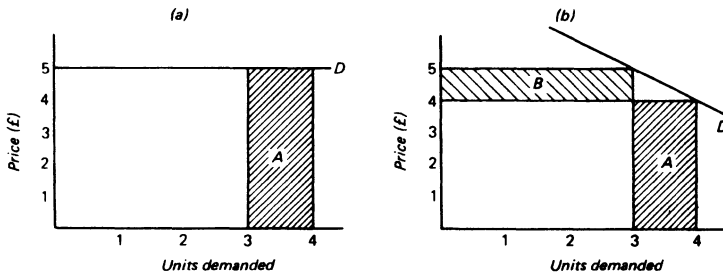

Figure 6.13 Marginal revenue under conditions of perfect and imperfect competition

In (b), however, the producer is selling under conditions of imperfect competition. If he wishes to sell a fourth unit, he must lower his price from £5 to £4. But this lower price applies not only to the fourth unit but also to the first three units. Thus his net addition to receipts is equal to what he gets for the fourth unit, *A*, less what he loses on the three previous units, *B*. Under imperfect competition, therefore, MR is always less than price at any given output.

(d) The relationship between the costs, revenue and output of a monopolist

Let us consider another imaginary manufacturer of lawnmowers, Airborne Mowers. To simplify, we shall assume that it has identical cost curves to those of Rollermowers, but differs in that it has a patent for its particular mower, thereby excluding competitors. In short, Airborne Mowers is a monopolist. Since its output is also the market supply, the number of the mowers which it puts on the market affects the price. Thus if it produces only 20 mowers a week, each will sell at £790; if total output is increased to 90, the price drops to £440.

Airborne Mowers has the same problem as Rollermowers – to decide which output yields maximum profit. But it has an extra complication on the revenue side – as output increases, price falls for the *whole* of the output. The result can be seen in marginal receipts (Table 6.4). These figures are plotted in Figure 6.14.

Table 6.4   Costs, receipts, and profits of Airborne Mowers (in £)

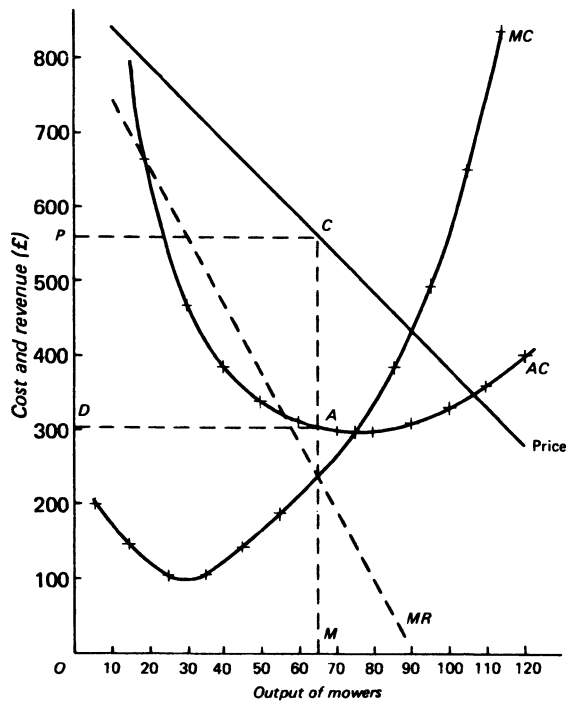| Output per week (units)	Total	Average total	Costs Marginal	Price per unit	Receipts Total	Receipts Marginal (p. unit)	Profits
0	10000	–	–	–	–	–	–10000
			200			840	
10	12000	1200		840	8400		–3600
			140			740	
20	13400	670		790	15800		2400
			100			640	
30	14400	480		740	22200		7800
			100			540	
40	15400	385		690	27600		12200
			135			440	
50	16750	335		640	32000		15250
			185			340	
60	18600	310		590	35400		16800
65	19825	305	240	565	36725	240	16900
70	21000	300		540	37800		16800
			300			140	
80	24000	300		490	39200		15200
			390			40	
90	27900	310		440	39600		11700
			510			–60	
100	33000	330		390	39000		6000
			660			–160	
110	39600	360		340	37400		–2200
			840			–260	
120	48000	400		290	34800		–13200

Figure 6.14 The equilibrium output of a monopolist

By inspection we can see that the maximum profit is made when 65 Airborne mowers are produced each week. At this output, MR = MC (both £240), as in perfect competition. But MR is no longer equal to, but is less than price (£565). Total weekly receipts are £36,725 and total costs £19,825 (by interpolation), giving a maximum profit of £16,900.

Alternatively we can use the price and ATC at an output of 65 units to calculate profit. In Figure 6.14, total receipts equal the rectangle *OMCP* (output times price) = 65 × £565; total cost equals the rectangle *OMAD* (output times average cost) = 65 × £305. Thus profit is the difference between the two: the rectangle *DACP* equals 65 × £260, i.e. £16,900.

(e) Monopoly and perfect competition: policy difficulties
Monopoly is an emotive word; it is often assumed that in seeking to maximise his profit the monopolist will always follow policies harmful to the consumer. The argument runs as follows.

Where there is perfect competition, output for all firms in the industry will take place where price equals MC, i.e. at *OM* (Figure 6.15). In other words, production is carried to the point, *OM,* where the cost of producing an extra unit, MP, just equals the value which consumers place on that extra unit in the market.
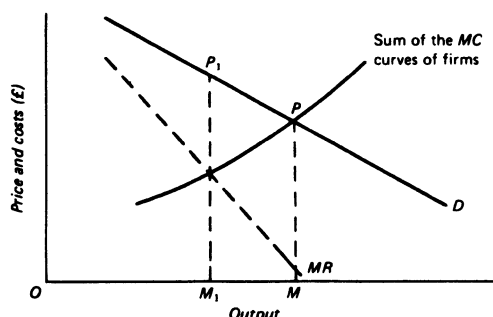
Figure 6.15 Output under perfect competition and monopoly

Now suppose a cartel formed from all the individual firms becomes responsible for selling the product. In order to maximise profits the cartel will sell an output where MC = MR, i.e. OM_1, at price P_1. Thus consumers get less of the product and at a higher price than under perfect competition. In short, factors of production are not fully allocated according to the wishes of consumers.

Yet, although it may contain much truth, the argument is not infallible.

First, it rests on the assumption that the competitive industry's supply curve will be the same as the MC curve of the monopolist. But this is unlikely to be so. A single firm may be able to obtain economies of scale not open to the comparatively small firms which comprise the competitive industry. In addition its investment may be higher, since it need not fear over-capitalisation of the industry through similar investment by rival firms.

It is probable, therefore, that the monopolist will, at the relevant market output, have lower costs than firms producing under perfectly competitive conditions. Indeed, even though the monopolist is maximising profits, the consumer may obtain more of the product and at a lower price than under perfect competition. Even so this does not achieve maximum economic efficiency for price is above MC, and super-normal profits are still being made.

Second, the argument ignores influences on the growth of firms over time, especially investment in research and innovation. Thus, we have to ask whether firms are more likely to spend on research and innovate if by being given monopoly powers they can be assured of the rewards. In short, is monopoly more conducive to growth than perfect competition? We cannot develop the argument here, but the mere existence of Patent Acts suggests that there is some truth in it. On the other hand there have been instances where monopolies have bought up patents so that they would *not* be developed in competition with them.

Third, if a monopolist can keep separate the parts of his market where elasticities of demand differ, he may, by charging different prices, supply certain clients who would be ruled out by a single price, e.g. by charging lower surgery fees to poor people, and giving students concessionary rail fares.

(f) The control of monopoly

The division of monopolies into 'spontaneous' and 'deliberate' does not make the first group 'white' or the second 'black'. In the first place our analysis has shown that no matter how the monopoly has arisen it will restrict output if its aim is to maximise profits. Second, while a monopoly may be 'deliberate' there may still be benefits from large-scale economies, etc. Thus, while some control of monopoly is usually desirable, it has to be applied according to the facts of the case, its benefits being weighed against its possible economic and social disadvantages – restriction of output, waste of resources in maintaining the monopoly position (e.g. by advertising), lack of enterprise through the absence of competition, exertion of political pressure to secure narrow ends (e.g. by trade unions), and a redistribution of wealth from consumers to the monopolist.

As a result monopolies in the UK are regulated rather than prohibited. Yet any policy faces difficulties. An exact assessment of the public benefits and disadvantages resulting from a monopoly is impossible. Very often, too, the decision as to whether a monopoly is useful or anti-social depends on circumstances and therefore varies from one period to another (note the fostering of monopolies in the depression of the 1930s). Moreover, if legislation is proposed the term 'unfair competition' has to be closely defined in rigid legal terms, whereas for purposes of control it really requires an elastic interpretation based on economic issues. Last, government policy in another field may affect a possible monopoly situation. Thus tariff protection, by restricting competition from abroad, fosters monopolies in the home market.

Broadly speaking, policy can take five main forms:

(i) *State ownership*

When it is important not to destroy the advantages of a monopoly, the state may take it over completely; the public then appears to be effectively protected. Freed from the objective of maximising profit, there should be no tendency for state-owned monopolies to seek high profits. Should, however, such profits be made, they would eventually be passed on to the public in lower prices, or in reduced taxation. In practice, however, lower profits may simply mask inefficiency in operation. Consequently provision must be made for the prices charged to be examined by an independent body and for efficiency checks to be carried out by independent experts.

(ii) *Legislation and administrative machinery to regulate monopolies*

This method is usually employed when it is desired to retain monopolies because of their benefits but to leave them under private ownership.

The Monopolies and Restrictive Practices Act 1948 (since amended) set up a Monopolies Commission to investigate monopoly situations. Upon the commission's report, a ministerial order could declare certain practices illegal. Subjects investigated have included the supply of electric lamps, household detergents, colour film, wallpaper, breakfast cereals, bricks, cross-channel ferries and perfumes; London rail services for commuters; collective discrimination; restrictive

practices in the professions; and proposed mergers, e.g. GEC and Plessey, Lonrho and the House of Fraser.

(iii) *Breaking up or prohibition of the monopoly*

Where the monopoly is, on balance, detrimental to consumers, policy can take the form of breaking it up or prohibiting it by legislation. Thus the state could reduce the period for which patents are granted or make their renewal more difficult. Alternatively it could outlaw attempts to eliminate competition, whether by unfair practices, the formation of cartels or restrictive agreements. Total prohibition was the policy at one time followed in the USA.

In the UK an investigation by the Monopolies Commission led to the Restrictive Trade Practices Act 1956. This (1) allowed manufacturers and traders to enforce *individual* resale price maintenance through the ordinary civil courts; (2) banned the *collective* enforcement of resale price maintenance through such practices as private courts, stop lists and boycotts; (3) required other restrictive pacts, such as common-price and level tendering, to be registered. For a practice to be allowed it must be justified as being 'in the public interest'.

But the 1956 act still permitted individual suppliers to enforce resale price maintenance for their own products. This was amended by the Resale Prices Act 1964, which made minimum-resale-price maintenance illegal, except for goods approved by the court. So far only minimum prices for books and proprietary medicines have been authorised.

The Monopolies and Mergers Act 1965 strengthened and extended the legislation on monopolies. A merger or proposed merger can be referred to the Monopolies Commission where it would lead to a monopoly or would increase the power of an existing monopoly. The act also increased the government's powers to enforce the findings of the Commission (for example, by allowing it to prohibit mergers or dissolve an undesirable monopoly).

The Fair Trading Act 1973 introduced a new concept with regard to monopoly and consumer protection. Unlike the earlier Monopolies Acts, whose primary concern was whether monopolies might be harmful to economic efficiency and thus not in the public interest, the object of this new act was stated to be to 'strengthen the machinery of *promoting competition*'. The Act:

   (i)   created an Office of Fair Trading under a Director-General with responsibility for discovering probable monopoly situations or uncompetitive practices. Thus the Office of Fair Trading provides ministers with information and advice on consumer protection, monopoly, mergers and restrictive practices.
   (ii) empowered the renamed Monopolies and Mergers Commission to investigate local as well as national monopolies and extended its powers of enquiry to the nationalised industries and even to restrictive labour practices (though with limited follow-up powers).
   (iii) reduced the criterion for a monopoly situation to a one-quarter (minimum) market share.

These powers were strengthened by the Competition Act 1980. The Director-General can investigate any business practice (in both the public and private sectors) which may restrict, distort or prevent competition. If found to be uncompetitive, he may accept an undertaking from the business responsible, or in default refer the practice to the Monopolies and Mergers Commission to establish whether it operates against the public interest.

In addition, EC rules providing for free and fair competition in trade between members enable the European Commission to ban restrive trade agreements and certain mergers.

(iv) *Price control*
The government can, by imposing a maximum price, restrict the monopolist's power over price.

(v) *Market solutions*

(1) FRANCHISING
To foster competition and to ensure that the state obtains a part of the profits from television, the right to broadcast independent television programmes over a given period is awarded mainly by the tender price submitted by each competing company.

(2) CONTESTABLE MARKETS
With the natural monopolies, where the initial fixed capital is so high that would-be new entrants could not accept the risk of irrecoverable sunk costs, it may be possible to get around this obstacle by allowing competitors to 'rent' the fixed capital. Thus independent gas suppliers have the right to negotiate contracts direct with large consumers and deliver supplies through the pipe-line system of British Gas. Even the threat of such competition may be sufficient to limit monopoly pricing by British Gas.

(3) PRICE REGULATION BY FORMULA
For certain newly-privatised monopolies, e.g. British Gas and British Telecom, price rises to certain consumers have been limited to the rise in Retail Prices Index, less 2 per cent (the assumed productivity increase). Any surplus revenue can be retained, thereby providing an incentive to achieve greater efficiency.

Pricing policy is also scrutinised by the respective regulating bodies, OFGAS and OFTEL.
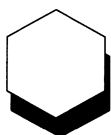
Part III
The environment

⑦ Externalities and cost-benefit analysis

7.1 Externalities

(a) Definition
In the pure market economy, resource allocation is the result of the decisions of consumers (households) and producers (firms) who seek to maximize the difference between benefits and incurred costs. We refer to these as *private benefits* and *private costs*.

But one weakness of the market economy is that it may fail to take account of any additional benefits or costs which 'spill over' from the original decisions. A firm may decide to build a new factory on a derelict site in a depressed district. In doing so it confers external benefits – tidying up the site and reducing the cost of unemployment benefit payments. On the other hand, should the factory be built in a residential district, it would incur spill-over costs of heavy vehicle movement, noise, loss of visual beauty, etc.

Externalities (spill-overs) are the costs or benefits additional to the private costs or benefits of a transaction and which are not provided for directly in the market price. Thus social costs (benefits) = private costs (benefits) + external costs (benefits).

(b) Externalities and the allocation of resources
The effect of externalities on the optimum allocation of resources can be shown diagrammatically (Figure 7.1). Suppose that a farmer applies nitrates to his field up to the point where the marginal revenue product (in terms of the value of the extra grass which will result) equals the marginal cost, OC kilos per hectare will be applied.

However, some of the nitrate may find its way into the water supply, and this increases as the application of nitrate per hectare increases.

This external cost has had to be added to the private cost (MC) to obtain the true marginal social cost, shown by the curve MSC, so that the socially efficient application of nitrates is reduced to OD per hectare.

7.2 Possible methods of dealing with externalities

It should not be assumed that the market economy completely fails to allow for externalities. Modern supermarkets, with their own extensive car parks, are partly a response to parking problems in city High Streets.

Or externalities may be provided for by private negotiation. Thus because weed cutting, although necessary, interferes with the fly-fishing on the
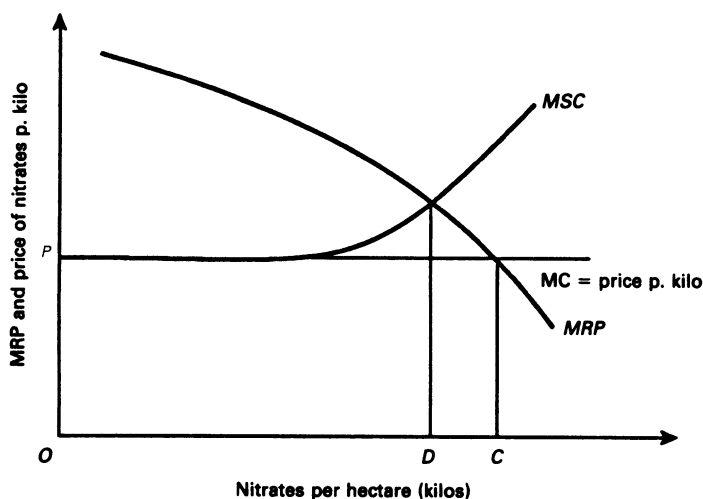
Figure 7.1   The external cost of nitrates applied to land

Hampshire chalk streams, owners have mutually agreed to restrict it to three specified weeks in the season.

Taking this a stage further, a pressure group may be formed, e.g. Greenpeace, to oppose what they perceive as an environmental cost of a proposal, or to protect an external benefit, e.g. the Worldwide Fund For Nature. Nevertheless private arrangements often prove inadequate.

Where 'free-riders' cannot be excluded, it may be impossible to organise sufficient collective bargaining strength to negotiate effectively. In any case, costs (or benefits) are often so far-ranging e.g. the detrimental effects of exhaust fumes, that not all the losers (or beneficiaries) can be identified. Usually, therefore, some form of government action is necessary.

Because there are a variety of methods by which externalities can be allowed for the government can choose according to the particular case.

First, it may introduce a pricing system to bring externalities into the reckoning. For example, to deal with traffic congestion, parking-meters may be installed.

Second, taxation and subsidies may take this a stage further. Thus the extra tax imposed on leaded petrol reflects its higher pollution effect. On the other hand, external benefits may be allowed for by subsidies, e.g. towards the costs of repairing ancient monuments and listed buildings.

Third, externalities may be covered by physical controls. Most evident are the planning consents required for building.

Fourth, externalities may be internalized by widening the area of control. The National Trust, for instance, harmonizes the interest, both of farmers and walkers, in order to secure maximum benefits from its Lake District properties.

Fifth, the government may itself assume responsibility for providing certain goods and services. This is usual when externalities are: (a) so extensive that only

government authority can adequately allow for them, e.g. providing a major airport; (b) cumulative, e.g., vaccination against small-pox.

To sum up, externalities are central to many of the problems facing governments today, and especially those concerning the environment. They therefore have to be incorporated into the decision-making process, where possible through the technique of cost-benefit analysis.

7.3 Why cost-benefit analysis (CBA)?

(a) The allocation of resources in the public sector

In the market economy, resources are allocated through the interaction of demand and supply in the market. Prices are the signals which coordinate the wishes of consumers with the cost of supplying goods.

But market signals may be either non-existent or defective. This applies particularly to many of the goods supplied by the government, e.g. roads, bridges, airports, parks, amenity land, education, health services, new urban areas and housing. Not only are such goods provided free or at less than cost, but their scale is such that externalities loom large.

Without firm market signals, decisions on the desirability of a project may rest mainly on subjective political views, e.g. subsidisation of city transport.

But allocating resources largely through the ballot box has serious defects in deciding investment in many public projects. First, the one-man, one-vote principle does not weight votes according to the intensity of satisfaction gained or lost. Thus, the simple majority decision could allow two voters marginally in favour of, say, a by-pass to outvote one who would suffer considerably from it. Second, economic efficiency in resource allocation requires that objective criteria should be used as far as possible.

(b) The nature of CBA

CBA is a technique which seeks to bring greater objectivity into decision-making. It does this by identifying all the relevant benefits and costs of a particular scheme and quantifying them in money terms to provide, as it were, a balance sheet upon which the ultimate decision can be made.

For example, the benefits of a new motorway would obviously include the time saved in travel, fuel economies, reduced congestion in towns through which motor traffic formerly passed, fewer road accidents and the pleasure derived by the extra motorists who could now make day trips. Against this, however, would have to be set the cost of constructing the motorway, the additional noise suffered by nearby residents, the congestion on the feeder roads, the toll of animal life and so on.

7.4 Difficulties of CBA

However, in giving a monetary value to such benefits and costs, we run up against both theoretical and practical difficulties.

First, there is likely to be some form of income redistribution. Thus those who suffer from the noise of the traffic on the by-pass, lose; the motorists and lorry drivers who save travelling time, gain. Only if the losers can be fully compensated by the gainers can we be satisfied that there has been no loss of satisfaction. But can we identify all those who are adversely affected by noise, and the extent of the noise on them? If not, full compensation cannot actually be paid.

Second, even if market prices are available, they may not, through indirect taxes, subsidies, imperfect competition or government price control reflect true opportunity costs. It is impossible, therefore, to achieve consistency in making adjustments.

Third, if no charges are made for the use of the by-pass, how do we know the possible value of the benefits received? The number of motorists likely to use the road can be estimated, but how do we value the journey each makes, since some are travelling on business and others on leisure pursuits? Similarly with the reduction in accidents; we can estimate the saving to the hospital service, in police time, etc., but how do we value the physical suffering avoided? And the by-pass may result in fewer deaths: what price do we put on the saving of human life? Similar problems arise in valuing such intangibles as noise, traffic congestion and the toll of animal life. It may be possible to obtain a price by analogy, e.g. the life-span earning-power of people dying in accidents, but no such calculation can be completely satisfactory.

Fourth, when estimating spillover costs and benefits which are not priced in the market, e.g. noise and human life, should we attach the same importance to them as for actual market prices when drawing up the balance sheet? Errors in such estimates, especially when 'shadow prices' form a large proportion of the balance sheet, could affect the viability of a project.

Fifth, there is the problem of deciding the cut-off point for externalities and the 'time horizon' for the benefits and costs to be included. The viability of a project could rest on these, and interested parties may be tempted to extend the cut-off point or the time-horizon to justify particular preferences.

Finally, in order to compare competing projects, we have to reduce them to present values by discounting future costs and benefits. But there are many different rates of interest to choose from, e.g. the rate at which the government can borrow, the rate for firms, the current rate determined largely by short-term monetary considerations. If there is less risk in a public project than in a private one, the lower government borrowing rate could be used. This lower rate is also supported by the fact that the government must consider the needs of future generations. Thus the social time preference rate is lower than the market rate since the latter is influenced by the limited time preferences of individuals.

7.5 An assessment of CBA

The theoretical and practical difficulties outlined above weaken CBA's effectiveness as a tool for decision-making. But its use is further limited by other considerations.

In particular, CBA cannot be used where political decisions dominate, e.g. the continued development of nuclear weapons.

Nor can a firm CBA decision be taken for a project which involves irreversible decisions, such as the survival of a species of animal or plant, since it is impossible to estimate the current cost of a decision which would deny future generations the opportunity to choose.

To sum up, while CBA provides a rational technique for appraising projects where market information is deficient, it must not make false claims for objectivity by dealing in precise sums. It is simply an aid to decision-making, not a substitute for it. Its main role, therefore, is to present systematically all the information relevant to a decision, indicating the weight which can be placed on the accuracy of the calculations submitted. Drawing up such an agenda ensures that all the relevant issues are fully debated before the ultimate political decision is taken.

⬡8 Protecting the environment

8.1 Economic aspects of the environment

(a) The environment as an economic good

The term 'environment' extends our earlier concept of land (p. 84) to include the sea and the atmosphere. As such it provides a flow of goods and services:

(a) materials (such as soil, minerals and timber) and energy (from fossil fuels, tides and wind);

(b) space, to produce food, erect buildings, develop communications and provide for sporting activities;

(c) the 'natural world', a consumer good in that it affords utility directly for walking, holidays, safaris, nature study or just enjoyment of peaceful surroundings;

(d) a 'sink' for waste products.

In essence, the natural environment is nature's capital. As such it is an economic good, being both 'scarce' and capable of being put to *competing* alternative uses.

(b) The role of economics

Since economics is concerned with the allocation of scarce resources between alternative uses, we have to ask what contribution it can make towards solving problems of the environment, especially those of conservation and pollution.

First, it can identify the major economic aspects, such as the link between a rising standard of living and the increasing demand of *future* generations for open spaces and buildings of historical interest. Second, it can indicate how externalities may be measured by CBA. Third, it can suggest appropriate policies, especially those that can operate through the market. Fourth, the economist must emphasise, not only the scarcity of the environment, but its indivisibility. Such resources as tropical rain forests, ancient woodland, fossil fuels, historical buildings and species of flora and fauna can be lost for ever as population increases and man seeks to produce more. And although in what follows the environment is discussed under the headings of 'conservation' and 'pollution', it must be emphasised that the problems are interlocking. Woodlands and forests may be conserved because of their richness in trees and flowers or because of their historical uniqueness. Yet their very existence fulfils the further function of absorbing the emissions of CO_2 from cars and power-stations and releasing oxygen. In other words, they are recycling waste and helping to prevent global warming. In con-

trast, the dumping of waste products, such as nuclear waste, cadmium and sulphurous gases can even threaten man's very existence.

Finally, economics should draw attention to the limitations of economic solutions, especially as regards the cultural and moral aspects. For instance, should it be left to the market to decide where a plant for processing poisonous substances is located? Local people may be so poor that they require the employment opportunities it can provide, and their economic weakness may force them to accept the accompanying health hazard.

8.2 Conservation

(a) The nature of conservation

With the exception of the special case of protecting an irreplaceable resource conservation is not simply preservation. Instead it seeks *creative continuity* by promoting vitality of use of the environment while ensuring that change is sympathetic to the quality of life for both present and future generations. Thus, in the UK, conservation embraces a wide field – green belts around towns, national parks, public bridle ways and footpaths, animal, fish, butterfly protection, sites of Special Scientific Interest, National Trust property, mineral and oil reserves, museums, buildings of special architectural and historical interest, and so on.

Certain aspects of conservation need emphasising. First, there is an *opportunity cost*. A 'green belt', for instance, keeps land in agriculture at a lower current market value than, say, housing, and also extends the journey to work of those city workers who live beyond it.

Second, externalities loom large. A farmer who rips out a hedge does so because a larger field will cost less to work. Yet there are external costs, e.g. for nature lovers and walkers who prefer the patchwork landscape of small fields.

Third, conservation is concerned with changes over time in both demand and supply. But, because individuals have a restricted time-horizon, it often means that decisions cannot be left entirely to market forces – particularly those concerned with estimating demand in the distant future, allowing for externalities, preserving stocks of renewable resources, and the impossibility of reversing wrong decisions.

(b) Conservation of stocks

Using a resource in excess of its capacity to reproduce itself eventually reduces supply to the point where its price rises, producing a contraction of demand.

With environmental resources which are renewable, e.g. fish, whales, grouse, conservation is mainly concerned with limiting what is harvested in order to maintain a stable stock.

With certain non-renewable resources e.g. fossil fuels and minerals, the rise in price as supply decreases, not only reduces demand but stimulates a search for substitutes.

But a special case of conservation occurs where resources, if lost, are irreplaceable, e.g. an historic building, ancient woodland, species of animals,

birds, insects, and flowers on the verge of extinction. Once they disappear, they are lost for ever. Here the over-riding objective is preservation.

In order to illustrate how economic analysis may assist in both forms of conservation, we take two examples – maintaining fish stocks and preserving an historic building.

(i) *Maintaining a renewable stock of fish*

Over-fishing means that over a period catches are so large that they exceed the rate of growth of the stock, which therefore declines in size. This could be a cumulative process resulting in the shrinking of stock below the minimum necessary for recovery.

The basic reason for the failure to conserve is the absence of private property rights over the fishing grounds. As a result, no one fisherman will voluntarily limit his catch, because that simply means that others can catch more.

To maintain the stock the annual catch has to be restricted.

Possible methods are:

(1) *Vesting fishing rights in a single body* which owns the fishing boats. Thus Iceland took the first step in this direction by confining fishing within 200 miles of its coast to Icelandic vessels. But this only goes a part of the way. Unless the Icelandic government also owns the fishing fleet, there must be some further control over the catch.

(2) *Imposing such physical controls* as: restricting the number of vessels by licence; enlarging the net mesh to ensure only mature fish of a certain size are taken; requiring each boat to stay in harbour for so many days a year.

   Not only are physical controls difficult to enforce, but they are resented by fishermen since marginal boats may be forced out of the industry.

(3) *Subsidising the de-commissioning* of fishing boats.

(4) *Taxing catches at so much a tonne* is administratively easier.

   Such a tax can also be flexible with different rates applied to different types of fish according to their relative scarcity.

(5) *Introducing quotas* for each ship, which in total will allow a specified aggregate catch. Since super-normal profit can be made by those fishing, these quotas could be sold to cover the administrative costs. Once in existence, quotas can be traded on a 'quota market'.

(ii) *Preserving a non-renewable resource: an historic building*

Not only is a decision to demolish a unique building *irreversible*, but it may be based on *defective market criteria*. If left to the market forces of demand and supply, the building would be demolished when its value becomes less than that of the cleared site plus the cost of rebuilding to the best alternative use, e.g. offices.

But *demand* criteria may be defective in that:

1. *external benefits* are not allowed for;
2. they ignore a possible '*option demand*' – people may be willing to pay something just to postpone a decision to demolish;

3. the present capital value of the building would be greater if, recognising society's longer time-horizon, the *social time preference rate of discount* were applied rather than the higher private time preference rate;
4. *rising future real income* may increase the demand for historic buildings (as with National Trust buildings).

On the other hand, an increase over future years in the *supply* of the alternative use (offices) would cause value to fall, with a consequent fall in the value of the cleared site. Thus, while the value of the historic building is likely to rise, the value of the cleared site declines. *Preventing immediate demolition* therefore allows for the possibility that eventually the value of the historic building could rise again above the value of the cleared site.

Government intervention can take different forms:

1. *Public ownership*, especially where the cost of excluding 'free-riders' is prohibitive, e.g. Hadrian's wall.
2. *Subsidising* the private owner through repair grants or tax concessions.
3. *'Listing'* the building, so that official consent is required for alteration or demolition. But this is only a 'stop-gap' measure until (4) can be agreed.
4. Permit the building to be *adapted to a more profitable use*, e.g. offices, while preserving its distinctive features. This avoids a charge on public funds.

8.3 Pollution

(a) Aspects of pollution

Pollution occurs when the introduction of waste matter causes damage to the environment. While pollution occurs through consumption (e.g. household waste, street litter), it is that resulting from production which is more serious (e.g. acid rain, smoke, gases, toxic chemicals, noise) for it can harm health, corrode buildings, damage fish, fauna and flora, and even threaten life on the whole planet through the 'greenhouse effect' of excessive carbon dioxide in the atmosphere.

Mainly it is the increase in industrialisation over the last century that has given rise to pollution. Yet it is the resulting prosperity which not only leads to the recognition of the *problem* of pollution, but can provide the means to combat it. For example, the EC's excess production of food has made it easier to care for the landscape.

Some unwanted residuals, such as carbon dioxide gas, can be transformed by the environment into harmless or even beneficial materials (e.g. oxygen). But this takes time and, as forests are converted into farmland, the CO_2 problem increases. We can therefore define pollution as occurring when the flow of residual waste exceeds the natural environment's capacity to absorb it.

(b) The economist's approach

In time technology may, while producing growth, contain pollution through such developments as: (i) substitute products which are more environmentally friendly, e.g. degradable containers; (ii) greater efficiency in production to reduce waste; (iii) on-site treatment of waste, e.g. desulphurisation of gases by power-

stations; (iv) the replacement of coal and oil with 'greener' sources of energy, e.g. natural gas, wind, tide.

However, such developments take time, and so what is 'sustainable pollution' must be assessed in the context of existing technology. The task of economics is to contribute to a solution of the problem by suggesting and examining a range of broad options.

In most cases pollution represents external costs. Thus in Figure 8.1, a chemical manufacturer would produce chemicals up to the point *OC* if there were no cost for discharging waste into the river. But spill-over costs give rise to a higher marginal social cost so that *OD* is the socially efficient level of production.

(c) Possible policies

Policy difficulties may arise because of the weaknesses of 'shadow' pricing in measuring intangible benefits lost, e.g. the natural beauty of the river bank. Moreover, since pollution occurs in different forms, circumstances and scale, there must be a range of policies so that the most relevant one can be chosen.

Possible policies are:

   (i) *'Greening' public opinion* to support waste recycling and energy saving.
  (ii) *Setting up an environmental protection agency* in which are vested *all* property rights, both private and external in order to bring externalities into private decision-making ('internalising externalities'). For example, the National Rivers Authority co-ordinates water supply, drainage, waste disposal and angling interests.
 (iii) *Negotiating a 'market price'* to avoid or reduce pollution. Thus Sweden assists Poland financially to reduce acid rain because this damages Sweden.
 (iv) *Imposing a maximum standard of pollution by regulation.* Here the government limits pollution by law, e.g. no discharge of oil waste by ships within so many miles of the coast. While this does allow the polluter to find the cheapest means of achieving the specified maximum, this standard tends to become the target, involves constant inspection and tends to impose national standards instead of allowing for local circumstances.

It should be noted that rigid control is essential where pollution is: (1) a threat to human life, e.g. asbestos dust, or (2) cumulative, becoming dangerous at a certain level, e.g. cadmium absorption by the soil.
  (v) *Subsidising the reduction of pollution* where it is impossible or too costly to identify the polluters (e.g. litter louts). Here the cost is borne by the taxpayer, not the polluter.

Alternatively, direct subsidies may be given towards: (1) the development of techniques to reduce pollution; (2) the production of cleaner substitutes, e.g. unleaded petrol; (3) the recycling of waste, e.g. bottles.
 (vi) *Taxing pollution* to ensure that the 'polluter pays' possibly according to the level of pollution (e.g. the extra tax on leaded petrol). Thus in Figure 8.1 a tax *EF* would reduce production to *OD*, where the level of pollution is acceptable. The major difficulty is that the polluter cannot always be identified or the pollution measured.
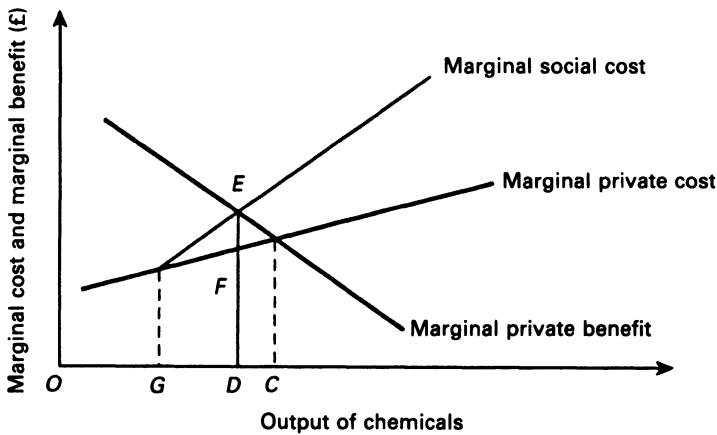
Figure 8.1   Efficient output with external costs

(vii) *Licensing pollution through tradeable permits*

Suppose in Figure 8.1 that output represents the aggregate of all chemical firms on a given river. The government decides to limit the discharge of polluting chemicals to *GD*. Each firm is given a licence to emit a share of *GD*. If the government wishes to raise revenue (equal, say to *EF*), it can sell or auction the licence. The essence of these pollution permits is that they can be traded on the 'permit' market. Those firms having a high cost of reducing emission will want to buy permits from the efficient firms who sell them for more than it costs them to abate.

Such selling of permits induces environmentally friendly producers to instal equipment to reduce pollution. At the same time, it uses the market to cover much of the regulation required. One difficulty is that as firms become more abatement-efficient, the supply of permits coming on to the market will increase, and their fall in price will allow inefficient firms to buy them. Here the government could itself buy on the market and, by confiscating permits, keep up the cost of pollution.

(viii) *International action.*

A 'sustainable earth' requires international agreement to deal with pollution e.g. from 'greenhouse' gases, acid rain and nuclear fall-out. Unfortunately national interests intrude, usually for economic reasons, e.g. whaling control. It really needs a United Nations organisation to be established as an 'environment protection agency'. Since it is the richest nations who have the greatest interest in preserving the environment, they should provide funds which can be used in different ways to compensate the adversely affected poorer countries, e.g. to preserve and maintain equatorial rainforests.
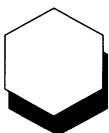
Part IV
Factors and their rewards

⑨ The determination of factor rewards

9.1 Introduction

(a) Sharing the national 'cake'

Factors of production co-operate together to produce the national product. Each of these factors is owned by somebody. How much of the cake each individual obtains depends upon (i) how much of the factors he owns, and (ii) the reward each factor receives.

Differences in individual incomes, therefore, depend upon both inequalities of ownership and inequalities in earnings. It is the latter which are the subject of this chapter.

(b) Factor rewards in a given industry, occupation or district.

Here we are concerned solely with the reward to factors in a given industry, occupation or district. That is, we examine how the price of a factor service is determined in a particular market. Analysis by ordinary demand and supply curves is therefore possible. (Later, [in Chapter 15], when we consider the economy as a whole, it is necessary to speak of labour, wages, capital, investment and the rate of interest in broad terms and substitute a general approach for this particular analysis.)

9.2 The theoretical determination of factor rewards

The theory which follows applies to all resources. However, it is usually illustrated in terms of labour and the wage-rate, and we shall adopt this practice.

The wage-rate is the price of labour and, like other prices, it is determined in a free market by demand and supply.

(a) Demand

The demand for labour is made up of the individual demands of all the firms using it. It is a *derived demand* – the factor is not wanted for its own sake but simply for its contribution to the product it makes. The actual price which a firm is willing to pay for a worker depends upon the addition to receipts which will result. We can be more precise by developing our analysis of the law of diminishing returns. Let us assume that (i) there is perfect competition in the market where the product is sold; (ii) there is perfect competition in buying labour – each firm is so small that it cannot, by varying its demand, alter the wage-rate which it has to pay; and (iii) in changing output, only the number of labourers employed is varied, other factors remaining fixed in supply.

The analysis of the law of diminishing returns (Chap. 6) was conducted in terms of physical yields – bags of potatoes. But when engaging labour, the firm is more interested in what the product sells for. What it asks, therefore, is how much will total receipts increase if an additional worker is employed? The value of this extra contribution is known as the *marginal revenue product* (MRP).

The MRP depends not only on the marginal physical product, but also on the price at which the product sells. Under perfect competition, the producer can sell any quantity at a given price. Hence the MRP is equal to the marginal physical product times the price of the product. Thus, in Table 6.1, by taking the marginal *physical* products and assuming that potatoes sell at £10 per bag, we can arrive at the MRP. For example, when 2 labourers are employed the marginal physical product is 14 bags, which at £10 a bag yields a total revenue of £140, and so on. (See p. 85.)

The farmer in our example will employ an extra worker so long as the MRP exceeds the cost, that is, the wage-rate. Thus, if the wage-rate were £130 per week, he would engage 6 workers because the value of the product of the sixth man was £130, and this just covered his wages. If fewer men, say 5, were employed, he could add more to receipts than to costs by taking on another worker, for the MRP of £150 would exceed the wage-rate (£130). On the other hand, if 7 men were employed, the farmer would be paying the seventh man £10 more than he contributed to receipts.

Of course it might be questioned whether the firm can always estimate the MRP of a factor of production. Thus, with certain workers, such as clerks, teachers and policemen, there is no definite physical product. How then can their marginal physical product, and thus the marginal revenue product, be measured? The answer is simply that it cannot be – but that does not alter the fact that, in practice, a firm behaving rationally and not 'empire-building' does proceed to engage workers as though it can so estimate.

The MRP at different wage-rates, therefore, gives the demand curve of the individual firm for labour (Figure 9.1). The *industry's demand curve* is the sum of the demands of the individual firms (curve *D*, Figure 9.2). This would be a simple horizontal addition if the price of the product remained unchanged. But it is much more realistic to assume that, as firms engage more labour, the extra output will lead to a fall in price of the product. The result will be that the industry's demand curve for a factor will fall more steeply than the curve obtained by a straightforward addition of firms' marginal-revenue-product curves.

(b) Supply
The supply of labour will depend upon:

(i) *The response of existing labour to a higher wage rate*
In the short period, an industry may find that a wage increase may result in *less* labour being supplied, the higher income enabling workers to enjoy more leisure, as in coal-mining. In the long period, however, higher wages should attract labour from other industries, occupations or localities, so that the long-period supply curve follows the shape of the *S* curve in Figure 9.2.
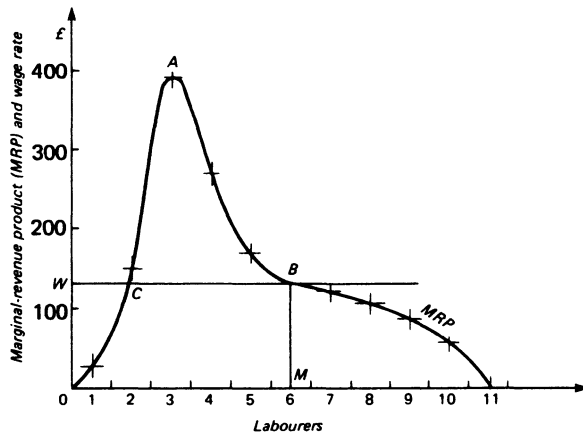
Figure 9.1   The firm's demand curve for labour

(ii) *The cost of attracting workers from alternative uses or localities*
In the long period, a higher wage will attract labour from other industries and occupations. The extent to which a given wage increase attracts workers depends upon the elasticity of demand for the products in these alternative sources. If demand is inelastic, higher wages can be offered to hold on to labour, and thus the supply of labour will expand little in response to the wage rise (see p. 101).

(iii) *The mobility of labour*
In the theoretical long period, a higher wage rate should attract labour from alternative uses or localities. But, because labour experiences particular difficulties in moving, the 'long period' is delayed indefinitely (see Chapter 10).

(c) Demand, supply, and the wage-rate
The reward of a factor, in this case the wage-rate, is determined by the interaction of demand and supply. Thus, in Figure 9.2, with demand curve *D* and supply curve *S* the wage rate is *OW*.

The wage-rate can rise through an increase in the MRP or a decrease in the supply of labour to the market. The MRP can rise through an increase in physical productivity or through a higher selling price of the product. Both would lead to a shift in the demand curve to the right, say from *D* to D_1 (Figure 9.2). As a result, the wage rate rises to OW_1 and the number of men employed increases from *OM* to OM_1. Higher labour productivity leads to higher employment, other things being equal. Similarly, a decrease in the supply of labour to a particular industry would have the effect of raising the wage-rate, but with fewer employed.

(d) The effect of imperfect competition on factor rewards
The above discussion assumed that there was perfect competition both in the sale of the product and in engaging factors of production. But in the real world such competition may not exist. If, for instance, the firm is selling its product under
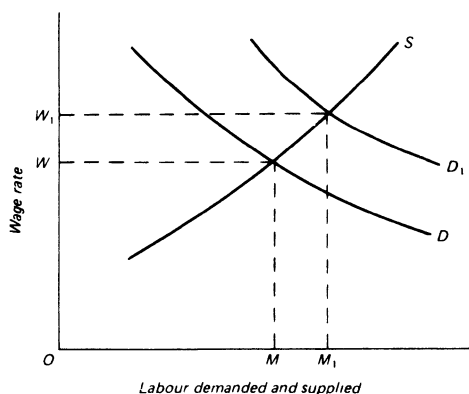
Figure 9.2   The determination of the price of a factor

imperfect competition, the price received will fall as output increases. This means that the marginal revenue product of an extra labourer will be less than the marginal physical product multiplied by its price, since the lower price applies to all previous units produced. Similarly, a firm may be the only employer of a factor in a locality, so that, as this firm demands more, the factor price rises not only for the additional factor but for all previous factors. Thus the marginal cost of employing such factors is higher than their market price. In both instances the demand for the factor will be less than it would have been had competition prevailed.

Finally trade unions may exert a monopoly power in the sale of labour, a subject discussed in Chapter 10.

⑩ Labour and wages

10.1 The nature of the labour force

(a) Special features of labour

Labour is the effort, both physical and mental, made by human beings in production. It is the 'human' element which is important.

Because people have feelings and emotions their response to economic forces is different from that of machines. First, whereas a machine which proves profitable can be reproduced fairly easily and quickly, the overall supply of labour does not depend upon its earnings. Other factors are more important in deciding how many children parents have (see p. 14). Second, the effort of labour is not determined solely by the reward offered. The method of payment may affect effort, while raising wages may result in less work being offered. Above all, a contented worker will produce more than an unhappy one; thus job satisfaction or loyalty to a firm, rather than a high rate of pay, may be decisive in inducing an employee to work overtime. Third, labour does not move readily, either occupationally or geographically, in response to the offer of a higher reward. Often such 'immobility' results from strong human contacts. Fourth, workers can combine together in trade unions. Finally, if unemployed for long periods, workers deteriorate physically and mentally.

Both firms and government must have policies which take account of these special characteristics. Training schemes are essential to improve the skill of workers and thus their productivity. Firms must pay particular attention to psychological and social factors in an effort to motivate workers, e.g. by profit-sharing schemes. Furthermore, they must endeavour to co-operate with the workers' trade-union representatives. Above all, firms have to comply with the constraints imposed by government policy (see p. 144).

This chapter examines these special features, though postponing the problem of full employment to Chapter 17–21.

(b) The supply of labour

The total supply of labour in an economy depends upon:

(i) *The size of population*

The size of the population sets an obvious limit to the total supply of labour. But while it is influenced by economic factors, e.g. through the birth rate and immigration, it is doubtful, especially in more advanced economies, whether economic factors are of paramount importance.

(ii) *The proportion of the population which works*
The working population, the proportion of the population which forms the labour force, is determined chiefly by age distribution, social institutions and customs, the participation rate of married women and the wages offered, and to a lesser extent by the numbers who can live on unearned incomes.

(iii) *The amount of work offered by each individual labourer*
Higher rates of pay usually induce a person to work overtime, the increased reward encouraging him to substitute work for leisure. But this is not always so. A higher wage-rate enables the worker to maintain his existing material standard of living with less work, and he may prefer extra leisure to more goods.

Nevertheless, as we shall see, more significant than the overall supply of labour are the obstacles to mobility which divide up the labour market.

10.2   The determination of the rate of pay

(a) Methods of rewarding labour
Some people are self-employed – window-cleaners, plumbers, solicitors, etc. As such they are really entrepreneurs, securing the rewards when demand is high but accepting the risks of working for a low return. Nevertheless, most workers contract out of risk, accepting a wage which is received whether or not the product of their labour is sold – although some element of risk-bearing may be incorporated in the wage agreement, e.g. by commission payments, bonus schemes, profit-sharing arrangements and the profit-related pay scheme introduced in 1986.

In what follows, reference will be mainly to the wage-rate – the sum of money which an employer contracts to pay a worker in return for services rendered. This definition includes salaries as well as wages, and makes no distinction between time- and piece-rates.

Earnings are what the worker actually receives in his pay-pocket (his 'take-home' pay) *plus* deductions which have been made for insurance, income tax, superannuation, etc. In practice earnings over a period often exceed the agreed wage-rate, additions being received for overtime working, piece-rates and bonus payments.

Where the nature of the work allows workers to be paid on a piece-rate basis as an alternative to time-rates, the firm has to consider their respective merits.

(i) *Time-rates*
Time rates are more satisfactory than piece-rate where:

(1)   a high quality of work is essential, e.g. in computer programming;
(2)   the work cannot be speeded up, e.g. in bus-driving and milking cows;
(3)   there is no standard type of work, e.g. in car repairs;
(4)   care has to be taken of delicate machinery, e.g. in hospital medical tests;
(5)   output cannot be easily measured, e.g. in teaching and nursing;
(6)   working long hours may undermine health, e.g. in laundry work;

(7)   the labour is by nature a fixed factor which has to be engaged whatever the output, e.g. secretarial and selling staff;

(8)   periods of temporary idleness necessarily occur, e.g. in repair work.

On the other hand, time-rates have certain disadvantages:

(1)   there is a lack of incentive for better workers;
(2)   supervision of workers is usually necessary;
(3)   agreements can be undermined by working to rule and 'go-slow' tactics.

(ii) *Piece-rates*

Where output is both measurable and more or less proportionate to the amount of effort expended, piece-rates are possible. It is not essential that each individual worker's output can be measured exactly. So long as the output of his group can be assessed he can share in the group's earnings.

The advantages of piece-rates are:

(1)   effort is stimulated;
(2)   the more efficient workers obtain higher rewards;
(3)   constant supervision and irksome time-keeping can be eliminated;
(4)   interest is added to dull, routine work;
(5)   workers can proceed at their own pace;
(6)   a team spirit is developed where workers operate in small groups;
(7)   workers are encouraged to suggest methods of improving production;
(8)   the employer's costing calculations are simplified;
(9)   output is increased, and the more intensive use of capital equipment spreads overheads.

We see, therefore, that piece-rates have advantages for both employee and employer. Moreover, the lower prices which result benefit the community as a whole. Nevertheless, for the following reasons they are often disliked by trade unions:

(1)   workers may over-exert themselves;
(2)   where piece-rates have to be varied according to local conditions or different circumstances, e.g. capital per employee, negotiations for a national wage-rate are difficult;
(3)   variations in piece-rates from one place to another undermine union solidarity;
(4)   the union may lose control over the supply of labour, and this makes it difficult to take strike action or to apportion work in periods of unemployment;
(5)   piece rates are subject to misunderstanding, e.g. an employer who instals a better machine may be accused of cutting the rate if he does not attribute all the increased output to the effort of labour;
(6)   workers may resist being shifted from tasks in which they have acquired dexterity (and which therefore produce high piece earnings) even though the current needs of the factory organisation require such a transfer. Thus

employers find that piece-rates lead to a loss of control over their employees, and many prefer to pay high time-rates to avoid this.

(iii) *Combined time- and piece-rates*
When deciding the basis of the wage-rate, both employees and employers want certain guarantees. Workers have a minimum standard of living to maintain, and they desire protection against variations in output which lie outside their control, e.g. weather conditions. On the other hand, employers providing expensive equipment must ensure that it is used for a minimum period of time. Thus piece-rates are usually incorporated in a wider contract which provides for some basic wage and a stipulated minimum number of hours.

(b) The wage-rate
In Chapter 9 it was shown how, under conditions of perfect competition, demand and supply determine the wage-rate. Take the wages of plasterers, for instance. The demand for plasterers depends upon the price at which houses sell (a derived demand) and the productivity of plasterers. The supply of plasterers is the number offering their services at different wage-rates. This will vary with the length of time under consideration. But in the long period more will be forthcoming the higher the wage-rate, since they will be attracted from lower-paid areas of occupations.

For example, if the conditions of demand and supply are different in different parts of the country, the wages of plasterers will differ. If there were perfect geographical mobility, plasterers would move from low-wage districts to high-wage districts, until eventually a common equilibrium wage-rate would be established. Similarly, where different wage-rates existed for different occupations, perfect occupational mobility would eventually eliminate these differences.

In practice, geographical and occupational mobility are not perfect, so that differences in wage-rates persist. A typist earns more in London than in Norwich; a doctor earns more than a docker. In short, immobilities result in the labour market's being divided into a number of separate smaller markets according to locality and occupation. It is necessary, therefore, to examine the obstacles to mobility in more detail.

(c) Obstacles to the mobility of labour
A worker may be required (i) to shift his job from one industry to another, (ii) to change his occupation, or (iii) to move his home to a different district. Often conditions dictate that all three types of change take place at the same time, but this is not necessarily so. Each presents its own obstacles to changing jobs, and gives rise to the 'immobility of labour'.

(i) *Obstacles between industries*
Provided that it does not involve a change of occupation or district, a worker can usually move his job from one industry to another fairly easily. Secretaries, storemen, lorry-drivers and porters, for example, are found in most industries. But middle-aged and older workers may experience difficulty. Prejudice or tradition

in certain industries may also prove to be obstacles. Women, for instance, would find it difficult to become taxi-drivers in London. Moreover, loyalty to a firm may prevent a worker from looking elsewhere even though he has suffered a cut in wages (though obviously this does not apply if he is made redundant).

(ii) *Obstacles to a change of occupation*
In changing occupations, obstacles may be encountered in both moving out of the old occupation or in entering a new one. They arise because:

(1)   a high natural ability is required in certain occupations;
(2)   training may be costly and take time;
(3)   stringent conditions are sometimes prescribed by trade unions and professional associations;
(4)   the new job may be repugnant; and, equally, some occupations, e.g. the church, art and acting, are so fulfilling that workers are not drawn into other occupations by the offer of a higher wage-rate;
(5)   workers may be too old to learn a new job;
(6)   workers may prefer to remain unemployed rather than accept a lower wage in an alternative occupation;
(7)   in spite of legislation, there is still discrimination on account of sex, colour, social class or religion;
(8)   workers may be ignorant of wage-rates and opportunities in other occupations.

Of the above, the greatest obstacle to occupational mobility is natural ability. In this respect it should be noted that there can be more mobility between occupations, e.g. storeman and clerk, requiring the same level of innate ability than between, for example, doctors and dockers, where there are marked differences in the natural ability and training required. The first is sometimes termed 'horizontal' occupational mobility; the second, where there are non-competing groups of workers, 'vertical' mobility.

(iii) *Obstacles to a change of district*
When it comes to moving from one part of the country to another, workers have to overcome both real and psychological obstacles. These include: the costs of moving; the difficulty of securing accommodation elsewhere on comparable terms, particularly for council and rent-controlled tenants; social ties involving friends, clubs, church, etc.; family ties, such as children's education; imperfect knowledge of vacancies or wages paid in other localities; and prejudice against certain parts of the country (people at present generally prefer to live in the south-east rather than in the industrial north).

Such immobility of labour means that wage-rates can often be more easily explained by supply conditions rather than by demand. Even if there is competition between employers, differences in supply produce differences in the wage-rates between occupations, and between localities even for the same occupation. Thus solicitors earn more than their clerks because, on the demand side, the services of solicitors are valued more highly and, on the supply side, the supply of

solicitors is small compared with clerks, for more natural ability and longer training are required.

Immobility is also one of the major causes of unemployment, and in Chapter 21 we consider some of the ways in which the government tries to reduce occupational and geographical immobility.

10.3   Trade unions and collective bargaining

(a) Objectives of trade-union activity

It would be wrong to regard trade unions primarily as a disruptive influence in the economy. For one thing there must be a means by which workers can communicate with employers. For another, by making the worker more contented, they enhance productivity. We can summarise their most important functions as: (i) improving working conditions; (ii) providing educational, social and legal benefits for members (iii) improving standards of work; (iv) obtaining pay increases; and (v) co-operating with governments in order to secure a workable economic policy and to improve working and living conditions generally. The remainder of this chapter is concerned with (iv).

(b) The process of collective bargaining

Collective bargaining is the settlement of conditions of employment by employers negotiating with the workers' trade unions. For its smooth working, certain conditions should be fulfilled. First, it must be pursued with good sense on both sides. This is enhanced where the industry has a tradition of good labour relations and where there is some accepted objective measure to which wage-rates can be linked (e.g. the Retail Prices Index, wage-rates paid in similar trades, the level of profits in the industry). Second, both sides should be represented by strong organisations. Where all employers are linked in an association, there is no fear of outsiders stealing a march by negotiating independent wage bargains, while, if the union can speak for all its members, employers know an agreement will be honoured. Unofficial stoppages damage the union's reputation, and to avoid them there must be regular contact between employer and union and prompt investigation of grievances on the shop-floor. Third, there must be an understood procedure for settling disputes. While this must not be so prolonged as to fray patience, it should exhaust all possibilities of reaching agreement before a strike or lock-out is called.

In short the procedure of collective bargaining covers (i) negotiation and (ii) the settlement of disputes.

(i) *Negotiation*

Broadly speaking the machinery for negotiation falls into three categories:

(1) VOLUNTARY NEGOTIATION

Generally the government has left it to the unions and employers' organisations to work out their own procedures, and today voluntary machinery covers nearly 70% of the insured workers of the UK. Because union organisation varies, the recognised procedure differs between industries and trades.

(2) JOINT INDUSTRIAL COUNCILS

Most industries have some national joint council or committee which, without out-side assistance, thrashes out agreements. Usually it follows the system of Joint Industrial Councils, composed of representatives of employers and workers in the industry. These consider regularly such matters as the better use of the practical knowledge and experience of the workpeople, general principles governing the conditions of employment, means of ensuring workers the greatest possible secu-rity of earnings and employment, methods of fixing and adjusting earnings, techni-cal education and training, industrial research, improvement of processes and proposed legislation affecting the industry. Although Joint Industrial Councils are sponsored by the government, they are not forced upon any industry, and some important industries, such as iron and steel, engineering, shipbuilding and cotton, which had already developed their own procedure for negotiation, have not formed Joint Industrial Councils. Moreover the tendency today is towards wage-rates being determined by local rather than by national agreements.

(3) THE AGRICULTURAL WAGES BOARD

With the abolition of Wages Council in 1993, the Agricultural Wages Board is the only remaining government-appointed body for fixing minimum wage-rates and related conditions.

(ii) *Settlement of disputes*

Where the negotiating machinery fails to produce an agreement, it is a help if agreed procedures exist for ending the deadlock. Three methods can be employed: conciliation, arbitration or special inquiry.

(1) CONCILIATION

In 1974 the Secretary of State for Employment set up an *independent* Advisory Conciliation and Arbitration Service (ACAS), controlled by a council whose members are experienced in industrial relations. When efforts to obtain settle-ment of a dispute through normal procedures have failed, ACAS can provide conciliation if the parties concerned agree.

(2) ARBITRATION

ACAS can, at the joint request of the parties to a dispute, appoint single arbitra-tors or boards of arbitration chosen from a register of people experienced in industrial relations to determine differences on the basis of agreed terms of reference.

Alternatively the Terms and Conditions of Employment Act 1959 allows claims that a particular employer is not observing the terms or conditions of employment established for the industry to be referred compulsorily to an industrial court for a legally binding award.

(3) INQUIRY AND INVESTIGATION

The Secretary of State for Employment has legal power to inquire into the causes and circumstances of any trade dispute and, if he thinks fit, to appoint a court of

inquiry with power to call for evidence. Such action, however, is chiefly a means of informing parliament and the public of the facts and causes of a major dispute, and is taken only when no agreed settlement seems possible.

The minister's power of inquiry also allows for less formal action, by way of setting up committees of investigation, when the public interest is not so general.

Neither a court of inquiry nor a committee of investigation is a conciliation or arbitration body, but both may make recommendations upon which a reasonable settlement of a dispute can be based.

(d) The extent to which trade unions can secure wage increases

This brings us to the questions of how and to what extent unions can secure increases in wage-rates by collective bargaining. We shall assume that the trade union is a 'closed shop' with 100% membership, making it virtually a monopolist in selling its particular type of labour.

Broadly speaking there are three ways in which a trade union can secure a wage increase:

(i) *It can support measures which will increase the demand for labour*

An increase in the demand for labour will come about if the MRP curve rises, either through an improvement in the physical productivity of the workers or through an increase in the price of the product.

The situation is illustrated in Figure 10.1. As marginal revenue productivity rises from MRP to MRP_1, wages of existing workers, ON, rise from OW to OW_1. Alternatively, if there were unemployment, extra men, NN_1, could be employed at the previous wage-rate.

(ii) *It can restrict the supply of labour, allowing members to compete freely in fixing remuneration with employers*

Where a trade union or professional association can limit entry it may also stipulate a minimum wage-rate or scale of charges. But it need not do so. The supply
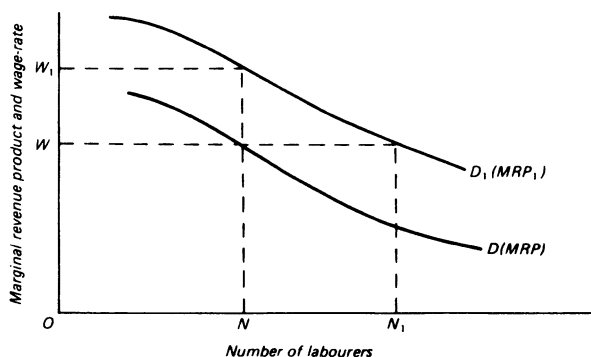
Figure 10.1 The effect on the wage-rate of a change in marginal revenue productivity
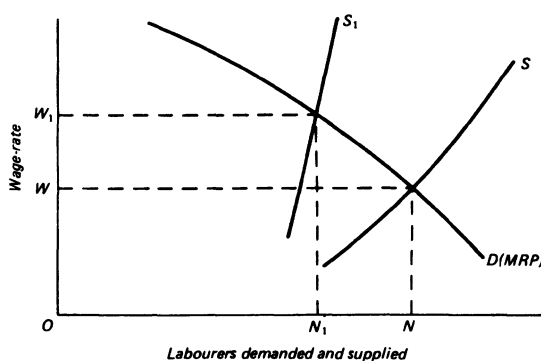
Figure 10.2 The effect on the wage-rate of trade-union restriction of the supply of labour

of plumbers and electricians, for instance, is restricted by apprenticeship regulations, but many work on their own account and *negotiate* their own rewards. Similarly, solicitors, doctors, surgeons, accountants and surveyors are restricted by the necessary professional qualifications, but suggested scale fees are not rigidly charged.

We can therefore analyse this method of securing a wage increase by the simple demand-and-supply approach (Figure 10.2). If the trade union reduces the supply of workers in an occupation from S to S_1, the wage-rate rises from OW to OW_1.

(iii) *It can fix a minimum wage-rate*
Where wages are raised by restricting entry, the trade union does not have to worry about unemployed members. It works simply on the principle that, assuming demand remains unchanged, greater scarcity leads to a higher reward.

Most trade unions, however, are faced with a more difficult problem. While they may secure higher wage-rates for their members, their success may be double-edged if, as a result, many members are sacked. What we really have to ask, therefore, is: *Under what conditions can a trade union obtain higher wages for its members without decreasing the numbers employed?*

Once again we have to consider conditions of competition.

(i) *Perfect competition in both selling the product and buying labour*
In the short period, even if there is perfect competition, a firm may be making super-normal profits. Here a strong trade union could, by threatening to withhold all its labour, force the employer to increase wages to the point where his super-normal profits disappear.

But this could not be permanent. The long-period equilibrium position is one in which there are no super-normal profits and the wage-rate is equal to the MRP. A higher wage will represent a rise in costs. Some employers will now be forced out of business (see p. 98) and remaining firms will have to

reduce their demand for labour until once again the wage-rate is equal to the MRP. Thus in Figure 10.2 we will assume that *OW* is the original wage-rate fixed by competition and *ON* the number of men employed – the trade-union membership. Suppose the trade union stipulates a minimum wage of OW_1. In the long period employment will then be reduced to ON_1. Given a downward-sloping MRP curve, this will always be true. Where there is perfect competition both in selling the product and in buying labour, a trade union can successfully negotiate an increase in wages only if there has been increased productivity: an increase without this will merely lead to members' becoming unemployed.

The amount of unemployment resulting from such a rise in wages depends upon the elasticity of demand for labour. This will vary with:

(1) THE PHYSICAL POSSIBILITY OF SUBSTITUTING ALTERNATIVE FACTORS

As the price of one factor rises, other factors become relatively cheaper and the tendency is to substitute them for the dearer factor. Thus if wages rise firms try to instal labour-saving machinery; that is, they replace labour by capital. But because different factors are imperfect substitutes for each other, such substitution is limited physically.

The degree to which substitution is possible is shown by the slope of the marginal productivity curve. Where labour is added to another factor, but is a poor substitute for it, marginal productivity falls steeply; where it is a fairly good substitute, marginal productivity falls more gently. Thus in Figure 10.3a labour is not a good substitute for the fixed-factor capital, and marginal revenue productivity falls steeply. Demand for labour is therefore inelastic, and a wage rise of WW_1 leads to only NN_1 men becoming unemployed. Compare this with Figure 10.3b, where labour and capital are better substitutes. Here the same wage rise leads to much more unemployment.

It should be noted that, since the possibility of substitution increases over time, the longer the period under consideration the greater will be the change in the labour force.
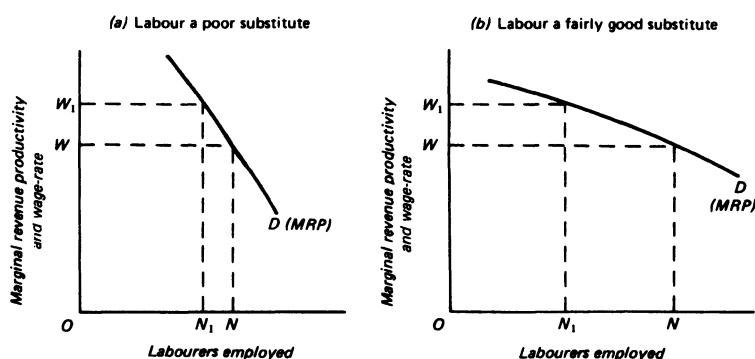
Figure 10.3   **Possibility of substitution between labour and capital**

(2) THE ELASTICITY OF SUPPLY OF ALTERNATIVE FACTORS

Under conditions of perfect competition the cost of a factor to an individual firm will not rise as the firm's demand for that factor increases (see p. 98). But when we are analysing a rise in the wage-rate of the workers of an *industry* we must recognise that the whole industry will now be demanding the alternative factors in order to substitute them for labour. This increased demand will raise the price of the alternative factors, again limiting the extent to which substitution is carried out. Thus if the supply of the alternative factor is perfectly elastic, only the physical considerations referred to above will affect the demand for it; if, on the other hand, supply is inelastic, it is likely that the quick rise in its price will soon make it uneconomic to substitute it for labour.

(3) THE PROPORTION OF LABOUR COSTS TO TOTAL COSTS

The proportion of labour costs to total costs has two effects. First, if labour costs form only a small percentage of total costs, demand for labour will tend to be inelastic, for there is less urgency in seeking substitutes (see p. 54). Second, if labour costs form a small percentage of total costs, as in steel production, a rise in wages will produce only a small movement to the left of the supply curve of the product. The opposite applies in each case.

(4) THE ELASTICITY OF DEMAND FOR THE FINAL PRODUCT

A rise in the wage-rate shifts the supply curve of the product to the left. Hence the market price of the good rises. We have to ask, therefore: How much will the demand for the good contract as a result? The answer depends upon the elasticity of demand.

If demand is elastic, D_{el} in Figure 10.4, the quantity of the good demanded will contract considerably, from OM to OM_1. This will mean a large reduction in the numbers employed. On the other hand, if demand is inelastic, D_{inel}, there will be no great contraction – only to OM_2. Here people are willing to pay a higher price for the good, OP_2, and this will cover the increase in wages. In other words, the MRP of labour has risen.
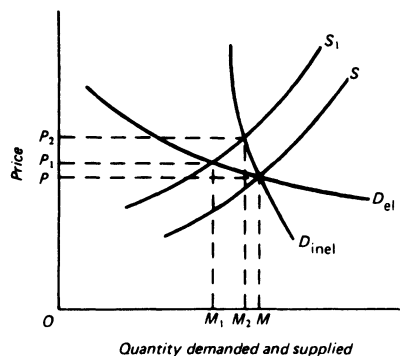
Figure 10.4 **The extent to which demand for the product contracts as a result of a wages increase**

Elasticity of demand depends mainly on the availability of substitutes. Thus demand in export markets is usually more elastic than in the home market, for with the former there are often many competing alternative sources of supply from foreign firms. Consequently, if an industry sells a high percentage of its output abroad, e.g. scientific instruments and machine-tools, the trade union is limited in its ability to secure a wage increase.

(ii) *Imperfect competition*

If there is imperfect competition in selling the product or in hiring labour, the firm is likely to be making super-normal profits. Here it may be possible for the trade union to wring increased wages from the employer without loss of employment. Since it is a monopolist in the supply of labour, the union can insist that the firm shall employ *all* or none of its members at the new wage-rate. Thus the firm may be forced to employ workers beyond the point where MRP = MC. The difference would come from super-normal profits, with the firm working on the principle that 'half a loaf is better than no bread'.

In these circumstances there is a whole range of possible wage-rates between the minimum which workers will accept and the maximum which employers are prepared to give rather than lose all their labour. The success of the trade union will depend, therefore, upon (1) the extent to which it can maintain its monopoly position by preventing employers from engaging blacklegs – non-union workers, or other substitute labour – and (2) the bargaining ability of its leaders relative to that of the employers. On the one side the union leaders have to estimate how high they can push the wage-rate without employers' allowing a strike to take place; on the other the employers must judge the lowest rate acceptable without a strike. As each is by no means certain of the other's strength, bluff will play a large part in the negotiations. Such factors as a large order-book for the firm's products, costly equipment's standing idle, a wealthy strike fund or increased profits, will obviously strengthen the union's hand. Considerations which could enlist public sympathy are a rise in the cost of living, a higher wage paid elsewhere in comparable occupations and an increase in productivity. Should a strike actually take place, it is usually because of misjudgment by one side; it is doubtful whether either really gains in the long run by strike action. Thus the strike is a form of 'blood-letting', allowing one or both sides to reassess the position before further negotiations take place.

(e) The government and wages

(i) *Influence on wage determination*

The government influences the wage-rate through: (i) its minimum-wage regulations for agriculture (see p. 139), (ii) the legal protection it affords to workers with regard to conditions of work, e.g. stipulating a written statement of the conditions of employment, prohibiting discrimination on account of sex or race, protecting employees against unfair dismissal, providing for redundancy payments and regulating conditions for health and safety at work; (iii) its efforts to break

down illogicalities, etc. sanctioned by custom; (iv) guidelines for wage settlements which it may lay down from time to time in its efforts to combat inflation.

(ii) *Curbing trade-union power*

With one main exception trade unions are similar to other pressure groups which seek to influence the government to further the interests of their particular causes. The exception is that trade unions can reinforce political means by economic sanctions. Moreover, in the case of key industries, just a small group of workers may, by strikes, go-slow tactics or working-to-rule, disrupt the whole economy.

The Thatcher government of 1979 took the view, therefore, that if it was to achieve its major role of stabilising the economy it could not allow its policies to be undermined by trade unions misusing their current extensive legal advantages. Thus it promoted a succession of legislation which progressively removed trade-union privileges.

In brief, the main changes effected are:

(a) Trade unions may be fined and their assets seized for offences committed by them or their members.

(b) Before industrial action can be taken, approval must be obtained by means of a secret ballot of its members.

(c) Picketing by employees must be confined to their place of work.

(d) Sympathetic strikes by workers not directly involved in the particular strike are illegal.

(e) A majority of members of a trade union must approve by secret ballot: (i) the setting up or continuance of a closed shop; (ii) any use of funds for any political purposes; (iii) the election of executive committees at least every five years.

⬡ 11 Capital, land and entrepreneurship

11.1 Capital

(a) What is capital?

A schoolteacher earns, say, £350 a week. She also has £1300 in the National Savings Bank, yielding her £52 per annum interest (£1 a week). We can say, therefore, that her total income is £351 a week, or £18,252 per annum; her capital assets are £1300.

Thus we see that, whereas *income* is a *flow of wealth* over a *period of time*, *capital* is a *stock of wealth* existing at any one *moment of time*.

This broad definition of capital, however, has slightly different meanings when used by different people. The ordinary individual, when speaking of his capital, would include his money assets, holdings of securities, and house, and possibly many durable goods, such as his car, television set, camcorder, etc. (sometimes referred to as 'consumer's capital'). The businessman would count not only his real assets (such as his factory, machinery, land and stocks of goods) but also any money reserves ('liquid capital') held in the bank, and titles to wealth (such as share certificates, tax-reserve certificates and government bonds).

But the economist considers capital chiefly as a form of wealth which contributes to production. In other words, he is concerned with capital as a *factor of production* – that is, as something real and not merely pieces of paper. It is the factory and machines, not the share certificates (the individual's entitlement to a part of the assets), which are vital to him.

This has two effects. First, in defining capital, he concentrates on producer goods and any stocks of finished consumer goods not yet in the hands of the final consumer. Second, in calculating the 'national capital', he has to be careful to avoid double-counting. Titles to capital – shares, bonds, savings certificates, National Savings, Treasury Bills and other government securities – must be excluded. Share certificates merely represent the factories, machinery etc., which have already been counted. Government debt refers to few real assets, for most has been expended on shell, ships, and aircraft in previous wars. The only exception regarding titles to wealth is where a share or bond is held by a foreign national, or conversely, where a British national holds a share or bond representing an asset in a foreign country. We then have to subtract the former and add the latter when calculating national capital. Foreign shares or bonds held by British nationals, for example, can always be sold to increase our real resources.

Naturally, 'social capital' (roads, schools, hospitals, municipal buildings, etc.) which belongs to the community at large is just as much capital as factories, offices, etc. And, in order to be consistent, owner-occupied houses have to be

WEALTH

INCOME   CAPITAL

Private individual   Economist

house   consumer's   money   securities
           capital   in bank   held

Businessman

tax-reserve   liquid   net debts   factory
certificates,   capital   outstanding   and
etc.   machinery

national capital

gold and   net assets in   owner-occupied   capital as a
foreign currency foreign countries   houses   factor of production
reserves

social
capital
                  producer goods   stocks of finished
                                        goods, including
                                        consumer goods
                                        not in hands of
                                        consumer

Figure 11.1 'Capital' in the economy

included, for they must be treated in the same way as houses owned by property companies.

(b) Capital as a factor of production

When the economist refers to 'capital', it is usually in the sense of *wealth which has been made by man for the production of further wealth*. This is because capital plays such an important part in increasing production and therefore in improving living standards. It is in this sense that the term is used from now on.

Increased production occurs because capital – tools, machines, irrigation works, communications, etc. – greatly assists people in their work. Indeed, with modern electronic equipment, machines often take over the work itself. As the use of capital increases, there are three possible gains. First, more current goods can be produced. Between 1984 and 1990 the output of agriculture, forestry and fishing increased by 12%, although the number of employees *fell* by 35%. There was thus an increased output per worker, due largely to more efficient machines and improved techniques resulting from capital investment in research. Second, instead of simply producing more current goods, people can be released to produce new goods. And, third, people can, as an alternative to more goods, enjoy increased leisure.

(c) The accumulation of capital

If capital is so important in adding to our well-being, why do we not have more of it? The answer is simply that we can accumulate capital only by postponing current consumption. In everyday language, more jam tomorrow means less jam today. The accumulation of capital represents an opportunity cost over time – consumption now or greater consumption later? A simple example will make this clear.

Suppose a peasant farmer has been tilling the ground with a primitive spade. By working 12 hours a day he can cultivate one hectare in a year. Obviously, if he had a plough which could be drawn by his oxen, it would help him considerably. How can be obtain it? Three ways are open to him:

(i) He could reduce the land he cultivates to $^3/_4$ hectare, using the 3 hours saved on tilling to make the plough.

(ii) He could reduce his leisure and sleeping time from 12 to 8 hours, using the extra 4 hours to make the plough.

(iii) He could decide not to consume some of the produce already harvested, exchanging it instead for the plough.

Whichever method is chosen entails some present sacrifice. With (i) and (iii) the farmer has less to eat, lowering his standard of living. With (ii) he has to go without some leisure. In short he has to draw in his belt or work harder. But the reward comes when he has the plough: with 12 hours' work a day he can now cultivate 2 hectares, thereby doubling his standard of living.

One other point emerges from this illustration: the more fertile his land, the easier it is for the farmer to increase his income. If, because of the poor type of soil, 16 hours were required to dig his 2 hectares, our farmer would have found it more difficult to obtain his plough. He could not reduce his consumption below the subsistence level; nor could he go without essential sleep. Similarly, countries with extremely low living-standards are in a vicious circle which can only be broken by economic aid from richer countries or by enforced five-year plans which ruthlessly cut current consumption, as in China and Cuba.

(d) Maintaining capital intact

Naturally our farmer will have to devote time to repairing the plough. So long as it is capable of cultivating 2 hectares, we can say that capital is being 'maintained intact'. If it is replaced by a better plough which allows more hectares to be cultivated, or another plough is added, capital is being 'accumulated'. Where capital is not maintained, it is being 'run down' or 'depreciated'.

In practice it is unusual for the same person to devote so much time to producing consumer goods and so much to the production of capital. Instead, production is organised by applying the principle of the division of labour – some people specialise in consumer goods and others in capital goods.

11.2 Interest

Investment, that is adding to capital goods or stocks, usually first involves obtaining liquid capital. Interest, expressed as a rate, is the price which has to

be paid for this liquid capital. What we shall examine here is the rate of interest which has to be paid for liquid capital in a *particular* use or industry. We shall *not* discuss what determines the *general* level of interest in the economy.

The *demand for liquid capital* arises because it is necessary or advantageous to use capital in production. The farmer who sows his seed in the autumn and harvests his crop in the summer is using capital in the form of seed. Similarly, a manufacturer needs capital in the form of a factory and machines because it is cheaper to produce in this way.

Now, as we saw when examining the farmer's decision to make a plough, the accumulation of capital can come about only by postponing present consumption. This can be done directly by the producer himself. The farmer could have obtained his seed by putting aside a part of the previous year's harvest; the manufacturer can buy his machines by retaining, rather than distributing, some of his profits. However, such retentions may be inadequate for the capital needed. In this case, funds might be borrowed from other people who have saved.

The actual demand of the farmer or manufacturer would depend upon the marginal revenue productivity of such capital and the rate of interest. For instance, if the addition to profit which a farmer thinks will be received during the year for adding an extra ton of fertiliser costing £100 is £120, he would be willing to borrow the money for the fertiliser so long as the interest he had to pay was not more than £20 – that is, 20%. The 'marginal-revenue-productivity-of-capital' curve will show the different amounts of capital which the farmer will find it profitable to borrow at different rates of interest.

The sum of the demand curves for liquid capital from all firms in an industry gives the demand curve for the industry, though some allowance should be made for a fall in the price of the good produced by the capital equipment (see p. 130).

The *supply* of liquid funds for a particular use can only be obtained by bidding them away from alternative uses. How much has to be paid for a given quantity relative to other uses will depend upon: (i) whether lenders consider more or less risk is involved; (ii) the period of the loan – people prefer to lend for a short period rather than a long one; and (iii) the elasticity of demand for other products employing capital.

Generally speaking, however, we can expect more liquid capital to be forthcoming the higher the rate of interest offered. We therefore have an upward-sloping supply curve, and the market rate of interest is fixed by the interaction of the demand and supply curves (as in Figure 9.2).

Once again, however, we must point out that this is only a partial explanation of the determination of a rate of interest. It does not tell us why, for instance, £10 million of liquid capital should be forthcoming at 10% rather than 6% or 14%. This will depend upon the general level of interest rates. Today the short-term rate at least is largely decided by the government in accordance with its major economic objectives.

11.3   Land and rent

(a) The determination of 'commercial rent'

To the economist, the terms 'land' and 'rent' have a special meaning. This is just as well, for in everyday speech each can imply different things. Thus, if I buy land for farming it will probably include buildings, fences, a water supply and a drainage system all of which are really capital. Similarly, I can rent things other than land – a house, a television set, a telephone, building equipment, shooting rights, etc. Rent in this sense simply means a periodic payment for the use of something. It can be termed 'commercial rent'.

Usually, however, rent does refer to payment for the use of a piece of land, and before we consider 'land' and 'rent' in their special economic sense, we must ask what determines how much rent is paid to a landlord.

The problem is similar to the determination of the return on any factor service. The demand for land depends on its marginal revenue productivity, and the curve slopes downwards from left to right for the reasons given in Chapter 9.2. On the supply side, land, like labour, can usually be put to alternative uses – building factories or houses, growing wheat or barley, raising cattle or sheep, and so on. A given piece of land will be transferred to its most profitable use. If, for instance, the price of cattle rises and that of wheat falls, some land will be transferred from arable to pasture farming. Thus the supply curve for land in a particular use slopes upwards from left to right. The interaction of the demand and supply curves will give the rent actually paid (as in Figure 9.2).

Of course, this assumes (i) that the landlord can vary the rent charged any time the demand for and supply of his particular type of land alters, and (ii) that land can be transferred fairly quickly to a different use. The first assumption is complicated by the fact that rents are usually fixed for a period of years. Only when the contract expires is the landlord free to adjust the rent. The second assumption implies that we are concerned only with the long period in our analysis. But what of the short period, when land is a fixed factor? An analysis of this situation is basically what we are concerned with in the remainder of this chapter.

(b) Ricardo's views on 'land' and 'rent'

To explain the special meaning which economists today give to the terms 'land' and 'rent', we have to examine the views of Ricardo, an early nineteenth century economist. He was concerned not with the rent paid to land for a particular purpose, but with the rent paid to land as a whole. Moreover, he was referring to land in the economist's sense – as the resources provided freely by nature (see Chapter 6) – and, as such, its total supply was fixed.

In this respect, he argued, land was different from the other factors of production, labour and capital, where more would be supplied the higher the price and, if no price at all were offered, there would be no supply. But with land as a whole – in the sense of space and natural resources – the same amount is available whatever the price offered. An increase in price cannot bring about an expansion of supply; and if the price fell to zero the same amount would still be available. Thus land as a whole has no supply price.

The return to land, therefore, was merely a 'residual' – the difference between, on the one hand, what was received for the product, and, on the other, the payments of wages and interest to labour and capital respectively. If the price received for the product was high, there would be more left over as rent; if the price was low, there would be less for rent. Rent did not determine the price of the good produced; instead, the opposite was true – rent was determined by price.

(c) The nature of the return to a fixed factor

Although there were certain blind spots in Ricardo's exposition of what determined price, he did point out an essential truth – that the return to a factor fixed in supply (that is, where supply is absolutely inelastic) will vary directly with variations in the price of the good produced by it. We can illustrate this more clearly by a simple example.

Let us assume: (i) a given plot of land on which only potatoes can be grown; (ii) that only land and labour are necessary to grow potatoes; and (iii) that the supply of labour for growing potatoes is perfectly elastic because only a small proportion of the total labour force is required.

The return to this plot of land will depend entirely on the price of potatoes. This can be seen from Figure 11.2. When the marginal revenue product of labour is shown by the curve QN, at a wage of OP, OM men are employed. The value of the total product is $OMNQ$; the wage-bill is $OMNP$ and the return of the plot of land PNQ. If now the price of potatoes increases, the marginal-revenue product of labour rises to Q_1N_1. OM_1 men are now employed at a wage-bill of OM_1N_1P. (Each worker still receives the same wage, OP, because the supply of this type of labour is perfectly elastic). But the return to the given plot of land has increased to PN_1Q_1. The opposite would apply if the price of potatoes fell.
Certain practical conclusions follow from the above analysis:

(i) Because the plot of land will grow only potatoes, it will be cultivated so long as the value of the total product is sufficient to pay the wage-bill. In
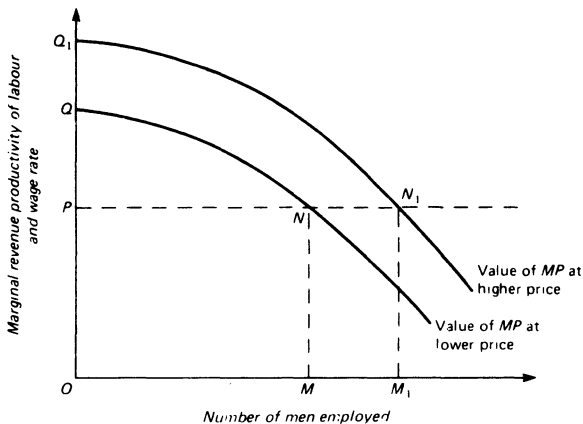
Figure 11.2 The effect of the price of a product on the rent of land

other words, at the lower price a lump-sum tax on the plot up to QPN
could be levied without affecting the output. This is the theoretical basis
of the often-proposed tax on land.

(ii)   The return to land as we have analysed it above – rent in its economic
sense – is purely a surplus. It arose because, by definition, our plot of land
was confined to one particular use – growing potatoes. The supply of this
land offered for sale or hire will not be affected by a price, simply
because nobody has any other use for it. In short, it has no opportunity or
transfer cost.

(iii) Once land has been built on, it is largely specific to a given use, and the
return to the land and building will be dependent on future demand.

(iv)   Because land is really space, it is impossible to increase the area of sites
in city centres except by building upwards, e.g. as in Oxford Street and
the City of London. Such fixity of supply means that rent is largely deter-
mined by current demand. Thus in Figure 11.3 it is assumed that the
supply of land is fixed at *OM*.

Demand for and supply of land
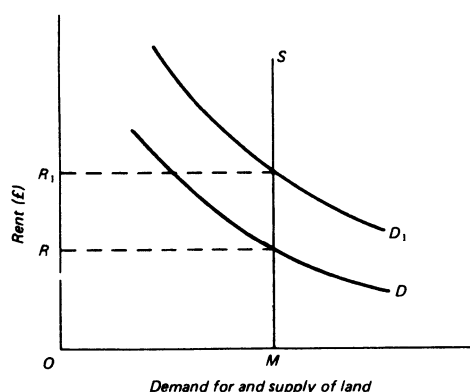
Figure 11.3   The determination of rent when land is fixed in supply

This means that the rent is determined by demand: an increase from *D*
to D_1 raises rent from *OR* to OR_1. For instance, rents in Oxford Street
depend upon the demand for shops there (which in its turn depends upon
people's spending) and rents in the City of London depend on the
demand for offices three (which in turn depend upon the level of business
activity).

(d) Economic rent

Economists have generalised Ricardo's concept of land to cover all factors which
are fixed in supply. 'Economic rent' is the term used to describe *the earnings of
any factor over and above its supply price*. Put in another way, it is any surplus
over its transfer earnings – what it could obtain in its next most profitable use (its

'opportunity cost', in our earlier terminology). How this idea can be applied generally will now be explained.

The actual rate of return to a factor is the price per period of time at which it is selling its services. For example, the return to a plasterer is his wage, say £360 per week. But what is the opportunity cost? Simply what has to be paid to retain it in its present use – that is, sufficient to keep it from going to the best alternative use. Take our plasterer, for instance. His next-best occupation may simply be plasterer's labourer, earning £180 per week. He would offer his services as a plasterer, therefore, at anything above £180 per week.

A second plasterer, however, may be a competent bricklayer, and as such earn £240 per week. He will only supply his services as a plasterer, therefore, if at least £240 per week is offered. And so we could go on. The supply curve of plasterers to the industry is thus an 'opportunity cost' curve (Figure 11.4).
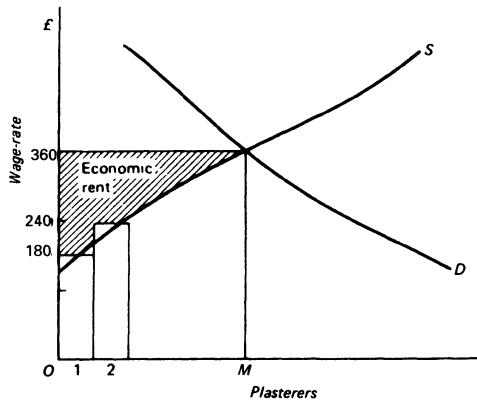
Figure 11.4   Economic rent

If in Figure 11.4 we now insert the demand curve, we can obtain the current wage-rate to the industry, £360, when *OM* plasterers will be employed. But all plasterers receive this wage-rate. Thus the first plasterer receives an economic rent of £180, the second £120, and so on. The total economic rent received by plasterers as a whole is shown by the shaded area.

(e)   What determines the size of economic rent?
The size of economic rent earned by a particular type of factor depends, on the one hand, on demand and, on the other, on the elasticity of supply of that factor and how the particular type of factor is defined.

(i)   *The elasticity of supply*
Elasticity of supply is determined largely by the period of time under consideration and by immobilities, some of which cannot be eliminated even in the long period. Both will affect economic rent.

Let us assume that, in the short period, the supply of plasterers is fixed: there is insufficient time for them to move into alternative occupations or for others to move in. In short, there is no alternative occupation – they can either work as plasterers or not at all. Thus all their earnings are economic rent – Figure 11.5a.
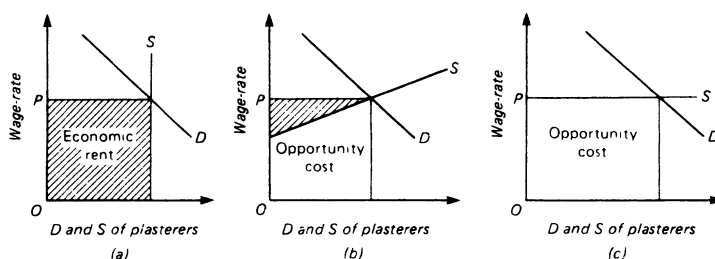
Figure 11.5   Economic rent and elasticity of supply

In the long period, however, other occupations can be trained as plasterers, and existing plasterers can move elsewhere. Sufficient has to be paid – the opportunity cost – to retain plasterers. Thus we have a long-period supply curve of plasterers, and economic rent is smaller (Figure 11.5b).

If the supply of plasterers became perfectly elastic, economic rent would disappear (Figure 11.5c). Thus economic rent depends upon a less than perfectly elastic supply curve to the industry.

Sometimes the degree of immobility between different uses or occupations persists indefinitely. Building sites for offices in the City of London, for instance, earn rents far in excess of what they could obtain in their best alternative use, say for houses. Simply because such sites are very limited in supply, competition for office accommodation has forced up the rents of these sites far beyond the possible price which houses could offer. A large part of their earnings, therefore, is 'economic rent'.

Occasionally, too, we refer to the 'rent of ability'. Many footballers, pop singers, film stars, barristers and surgeons have talents which, to all intents and purposes, are unique, for they cannot be duplicated by training others. Their high earnings, therefore, are almost wholly in the nature of 'economic rent'.

(ii) *The definition of an 'occupation', etc.*
If we adopt a wide definition of our factor, e.g. land as a whole, the distinction is between employing it or idleness, and thus the whole of its earnings is economic rent. This is what Ricardo had in mind.

If, however, our definition is narrower, e.g. covering only land for a particular use, such as growing wheat, then the opportunity cost (e.g. growing barley) will be larger and the economic rent smaller. Similarly, we could distinguish between cabinet-makers and carpenters, surgeons and doctors, and so on. Each would give a smaller 'economic rent' than if the distinction were simply between cabinet-makers and labourers, surgeons and nurses, etc.

(f) Quasi-rent

For fixed factors, particularly capital equipment, what the firm has to pay to retain them will vary according to the period of time.

In the short period, capital equipment is, by definition, fixed in supply. There is no transfer price. More capital equipment cannot be added; nor can existing equipment be diminished. The firm will continue to work its capital equipment so long as total earnings just cover the cost of variable factors (see Chapter 6). Any earnings above variable costs will be in the nature of a residual which helps towards the cost of the fixed factors. The size of this residual depends upon the price at which the product sells.

This can be seen immediately if we refer back to Figure 6.3. Were the demand for mowers to increase, the price would rise, say to £600, and production would be expanded to the point where once again price equalled marginal cost – that is, to 100 units. The increased cost of such production would be equal to the increase in total cost – that is, £5,100. But total receipts would have increased by £19500, and so the fixed factors earn an additional return – the increase in 'super-normal profit', equal to £14,400.

As time passes, however, we move into the long period. If the product has been selling at a high price, the high return to the capital equipment will induce firms to instal additional equipment. On the other hand, if the price of the product was low, existing capital equipment will either be transferred to its next most profitable use or, when it wears out, simply not be replaced. In the long period, therefore, earnings of fixed factors are, under perfect competition, equal to their transfer cost: economic rent is eliminated.

To distinguish between economic rent which is more or less permanent and that which disappears over time, the latter is often referred to as 'quasi-rent'. It is not a true rent, for the high return earned by such factors leads to an increase in their supply, and this eliminates the economic rent they earn. True rent refers only to factors which are fixed in supply; even if their earnings are high, identical factors are not forthcoming, and so economic rent persists.

11.4　Entrepreneurship

(a) The identity of entrepreneurship and risk-bearing

For production to take place, resources must be brought together and set to work. Whoever undertakes this task is often described as 'the entrepreneur'. Usually, however, economists give a narrower meaning to the term.

Organising production can be broken down into two parts. First, there is the task of co-ordination – setting the resources to work. Second, there is accepting the risk of buying factors to produce goods which will only be sold in the future – when receipts may not cover costs.

In practice it is not always easy to separate co-ordination and risk. A farmer, for instance, not only manages and runs his farm, but also accepts the risk involved in deciding what to produce. On the other hand, in a joint-stock company most co-ordination is left to a paid board of directors, with a manager under them. Here the risks of the business are borne by the ordinary shareholders; with

a public corporation, they are carried by the taxpayers. Neither take a part in running the business, except remotely.

The function of co-ordination, therefore, can be fulfilled by a paid manager. In this respect, management is simply an exceptionally highly skilled form of labour. Thus we narrow our concept of enterprise to 'bearing those risks of the business associated with ownership'.

(b) The nature of risks

A business is always open to the risk of fire, accidents, burglary, storm damage, etc. But these risks are calculable. A mathematician can work out, for instance, the chances of a building's catching fire during the year. While he cannot say which building will be destroyed, he knows that on average, say, one out of every 10,000 will be. Such risks, therefore, can be insured against. They are then reduced to a normal cost – what the firm has to pay to contract out of the risk involved.

Certain risks, however, cannot be calculated according to the law of averages. Nobody, for instance, can forecast with certainty how many cold drinks will be sold in Britain next summer. That will depend on the weather. Similarly, it might be thought that a new 'mini' car will sell well. But again, there is a chance that it will not. The risk that demand will be different from that estimated cannot be reduced to a mathematical probability. Such a risk, therefore, cannot be insured against; it must be accepted by those persons whose money is tied up in producing goods for an uncertain demand.

These uninsurable risks are inherent in a dynamic economy. Modern methods of production take time. When an entrepreneur engages resources, therefore, it is an act of faith – faith in his estimate of the demand for the product some time ahead. But demand can never be completely certain. People have freedom of choice, and their tastes may change. Many of the factors affecting demand fluctuate even over a relatively short period of time. It is similar on the supply side. Techniques do not stand still; new methods discovered by a rival may mean that, by the time a firm's product comes on the market, it is undersold by a cheaper or better substitute.

Thus there is always some degree of uncertainty, and this involves risk. It is a risk which must be shouldered by those who back with their money the decision as to what shall be produced. The true entrepreneurs, therefore, are those who accept the risks of *uncertainty-bearing*.

11.5   Profit

(a) How profit differs in nature from other rewards

The reward of uncertainty-bearing is 'profit'. But profit differs from the earnings of other factors. First, it may be negative. Whereas wages, rent and interest are paid as part of a contract at the time of hiring, profits are received in the future, and then only if expected demand materialises. Thus its size is uncertain. Where the entrepreneur has been far too optimistic, a loss is made. Second, profit fluctuates more than other rewards, for it feels the immediate impact of booms and

slumps. In a boom, profits rise faster than wages, while in a slump they fall more severely.

(b) Differences in the meaning of the term 'profit'
It is essential to distinguish four different concepts of 'profit'.

(i) *Profit in its everyday meaning*
To the accountant, profit means simply the difference between total receipts and total costs (see p. 88). But, because the economist defines cost in terms of alternatives forgone, he would amend this idea of profit by deducting, first, the return which would have been received on capital had it been used elsewhere, and second, the value of the entrepreneur's skill in the best alternative line of business (see also p. 88).

(ii) *Normal profit under perfect competition*
Because uncertainty cannot be eliminated from a dynamic economy, there must be a return to induce people to bear uncertainty. This is true even in the long period. Thus there must be a rate of profit – the price which equates the demand for and supply of entrepreneurship. In the long period under perfect competition, any rent element from profit is eliminated. We then have normal profit – the cost which has to be met if the supply of uncertainty-bearing is to be maintained.

Two modifications should be noted. First, industries differ as regards the uncertainty involved. Where fashions or techniques change frequently, for instance, uncertainty is greater. Entrepreneurs in such riskier enterprises, therefore, require a higher level of normal profit. Second, the elimination of the rent element in profit in the long period is only possible if one assumes that entrepreneurs of equal ability are available. In practice this is not so. Thus there will always be some entrepreneurs earning a rent of ability (super-normal profit) even in the long period, simply because their forecast and decisions are more accurate.

(iii) *Super-normal profit*
Under perfect competition the entrepreneur is able to make super-normal profit for a period because new firms cannot enter the industry. Certain factors such as key workers and machines are for a time fixed in supply, and entrepreneurs already possessing them will make super-normal profit. In others words such profit is really the return to fixed factors in the short period; it is the 'quasi-rent' earned by such factors.

(iv) *Monopoly profit*
With monopoly, competitors can be excluded. Certain factors, e.g. diamond-mines, know-how, patents and copyrights, are fixed to the monopolist. Even in the long period, competitors cannot engage such factors, and so super-normal profits persist. The profits of the monopolist, therefore, are closer to economic rent than to quasi-rent.

(c) **The role of profit in the market economy**
'Profit' tends to be an emotive word, and firms which make large profits are often frowned upon. But usually there is little justification for this, since it is through profits – and losses – that the market economy works. We must emphasise, however, that we are discussing only profits under competitive conditions. But, given such conditions, profit fulfils the following functions:

(i) *Normal profit induces people to accept the risks of uncertainty*
Because uncertainty is implicit in a dynamic economy, a reward – normal profit – is essential for entrepreneurs to undertake production. Thus normal profit is a cost, as essential as the payment of wages. The level will vary with the industry; thus it will be higher for oil exploration than for selling petrol.

(ii) *Super-normal profit indicates whether an industry should expand or contract*
When a firm produces a good which proves to be popular with consumers, it probably makes super-normal profit. This indicates that output should be expanded. On the other hand losses show that consumers do not want the good, and production should contract.

(iii) *Super-normal profit encourages firms to increase production*
Profits not only indicate that consumers want more of a good; they are also the inducement to firms to produce those goods. As we saw in chapter 6, super-normal profits act as the spur for existing firms to increase capacity and for other firms to enter the industry. On the other hand, when losses are being incurred, firms go out of production and the industry contracts. Thus losses are as important as profits in the operation of the market economy.

(iv) *Super-normal profit provides the resources for expansion*
An industry making super-normal profit can secure the factors necessary to expand. First, profits can be ploughed back, while shareholders will respond to requests for further capital, usually through rights issues. New firms can enter the industry, because investors will subscribe to a company intending to operate where the level of profits is relatively high. Second, profits allow expanding firms to offer higher rewards to attract factors. In this way resources are moved according to the wishes of consumers.

(v) *Super-normal profit encourages research, innovation and exploration*
Research, e.g. for new drugs, and exploration, e.g. for oil, carry a high risk of failure and therefore of wasted capital expenditure. But the possibility of high returns if successful induces firms to engage in research, especially if new developments are protected for a period from competitors by patents.

(vi) *Profits ensure that production is carried on by the most efficient firms*
In a competitive industry the firm making the largest profit is the one whose costs are lowest. It will have an incentive to expand production and, if necessary, can afford to pay more for factors to do so. Less efficient firms must copy its methods

to retain factors. In any case the increased output of the more efficient firm will eventually lower the price of the product. As a result inefficient firms make losses: profits have become negative.

To sum up, profits and losses are the means by which the process of natural selection occurs in the market economy. Where there is competition it is wrong to regard profits as being somehow immoral. The exception is monopoly profits, which are not eliminated even in the long period. Entry into the industry is not free; consequently profits are not competed away. It may be that such monopoly profits stimulate research and allow an industry to expand. But where scarcity has been deliberately brought about, they simply represent an economic rent earned at the expense of consumers by the monopolist owners. Thus an efficient allocation of resources according to the wishes of consumers does not take place.
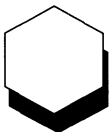
Part V
Public-sector production

⟨12⟩ The provision of goods and services through the public sector

12.1 The case for public-sector production

(a) Defects of the market economy in allocating resources
In the market economy, supply responds to price signals. The main strengths of this method of allocating resources are (i) that, through demand, individual preferences are indicated by their effect on the price of a good; and (ii) that, through the profit motive, resources are used efficiently in supplying goods. Chapters 3–6 have explained how both are achieved, and how some inefficiencies, e.g. those of monopoly power, may be corrected by government action.

In certain circumstances, however, such market intervention is inadequate. First, *community goods*, such as defence and street-lighting, cannot be supplied through the price system since it is impossible to exclude free-riders. Second, *public goods,* such as parks and roads, where there is no reduction in the quantity available for others when one person has more (that is, marginal cost is nil), are not fully used when charges are levied. Third, either through ignorance or miscalculation of future benefits, people – even those with sufficient income – may devote an inadequate proportion of it to purchasing *merit goods*, such as education. Fourth, a project may be so large that only the government can provide the *capital* required, particularly where there is some doubt if revenue will cover total *costs,* e.g. Concorde, atomic energy. Fifth, *external costs and benefits* may be so widespread that only the government can take full account of them, e.g. new-town development, new-airport construction, inner city and dockland regeneration.

(b) Forms of public-sector organisation
Where any of the above conditions hold, there is a case for providing the good or service through the public sector. The actual form of organisation will vary between a government department, a local authority, a nationalised industry and other quasi-government bodies – usually termed commissions, committees or councils.

In 1994 the public sector accounts for 25% of both Gross Domestic Product and the labour force. This chapter considers some of the broad problems concerning the public sector and then examines the efficiency of each form of organisation in the context of the special objectives of the services provided.

12.2 Demand and needs

(a) Differences between 'demand' and 'needs'

Whereas goods and services are supplied by private-sector firms in response to effective demand, government departments and local authorities provide goods and services according to 'needs', a social rather than an economic concept. 'Needs' are more difficult to assess than demand.

For example, in the private sector owner-occupied houses are built according to the price which people are able and willing to pay for them. Demand will depend upon the price of the house, the prices of other goods and services (particularly near-substitutes), the level of income, the distribution of wealth and all the other factors mentioned in Chapter 3 as influencing the conditions of demand. Supply responds automatically to this demand; the number and type of houses supplied depends ultimately on the equilibrium price determined in the market.

In contrast, in providing housing according to needs the public-sector authorities regard housing as a social obligation. Consequently, price signals are either inadequate or non-existent. This increases the difficulties of decision-making. Consider the factors which have to be borne in mind in planning a housing programme based on needs. First, the authorities have to estimate the number of households seeking accommodation according to the sizes of the family units, the ages of their members, their location, their preferences as between houses and high-rise flats, and so on. Moreover, since houses are very durable, some consideration has to be given to future requirements. Second, the authorities have to decide arbitrarily on the standard of an adequate housing unit. Third, they have to get the dwellings built, either through a private contractor or by their own direct-labour building organisations.

(b) Subjective assessment of 'needs'

The task of estimating needs is made more difficult because there is no price system in operation to provide reliable criteria. Thus rents charged by local authorities are less than the open-market rent. This means that demand exceeds supply, and the only indication of need thrown up by this restricted-price system is the number of households waiting their turn on the housing list.

And, all the time, the authorities must be conscious of dealing with limited resources – more spent on roads may mean less available for the health services. In the last resort, therefore, the standard of goods and services provided on the basis of needs is determined by the political views of the central government and local councils.

12.3 Pricing policy in the public sector

The problem arises as to how goods and services provided by the public sector are to be paid for.

(a) Provision from the proceeds of taxation

With *community goods,* where free-riders cannot be excluded, no price can be charged, since nobody will pay when private rights to them cannot be granted, e.g. with defence, consular services and street-lighting. Here the cost has to be covered entirely from taxation.

The same largely applies to *public goods,* where in economic terms marginal cost is nil (see p. 163). These will only be enjoyed to the maximum if no charge is levied, e.g. for crossing bridges, visiting public parks, using motorways, etc.

(b) User-charges

With other goods and services, some charges can be levied. Indeed such charges promote economy in use, while there is an element of fairness in that those enjoying them contribute to their cost. Nevertheless, *merit goods* are usually subsidised from taxation, mainly in order to encourage people to consume more than they otherwise would if they had to pay the full economic price, e.g. for housing. Redistribution of income, however, may also be an objective of such subsidies, e.g. housing subsidies benefit low-income families.

With some services technical considerations may be decisive. Thus while motorways could be financed by toll charges, the effect on the traffic flow, especially during vital rush hours, has so far led the UK to pay for them from general taxation. However, users contribute heavily through motor vehicle licences and petrol duties, while charges are imposed on motorists in minor matters, e.g. parking fees.

The real difficulty for pricing policy arises with those goods where total costs are expected to be covered by charges on consumers. This applied particularly to the nationalised industries, which were generally expected to pay their way. The snag is that if, to secure the optimum allocation of resources, production takes place where price equals marginal cost (the principle of marginal-cost pricing), total costs may not be covered. The reason is that because initial fixed costs are proportionately so high, e.g. railways, a single price equalling marginal cost cannot generate sufficient revenue to cover total costs.

In practice the problem has been overcome in three ways:

(i) The difference has been covered by a *subsidy,* either directly, e.g. for city transport, or indirectly, through writing off accumulated deficits from time to time, e.g. for coal and railways.

(ii) A *standing charge* is levied irrespective of units consumed, e.g. for electricity. The standing charge goes to meet fixed costs; the price per unit consumed covers marginal costs.

(iii) The industry is allowed to exploit its monopoly position by *price discrimination.* This is possible where different customers, having a different elasticity of demand for the product, can be kept separate, each being charged the price he is willing to pay. By 'charging what the traffic will bear', total revenue is increased. Such price discrimination by consumer category is used by British Rail where, for example, cheap-day trippers, senior citizens and students are charged lower fares than commuters.

12.4   The government department

(a) The need for 'public accountability'

When the state provides goods and services, consideration has to be given to two fundamental principles which pull in opposite directions.

The first, 'public accountability', arises because British democracy requires that where the state is granted powers it shall be answerable, in some form or another, for the way in which those powers are exercised. The citizen requires some assurance that powers granted to the state to produce goods and services are not abused by authoritarianism, inefficiency or monopolistic exploitation.

The second principle is 'economic efficiency'. The difficulty is that, by insisting on strict public accountability, we may so tie the hands of those running the state services that they cannot operate efficiently.

(b) Accountability through the government department

The highest form of accountability is achieved when a government department produces the goods or services, for at its head is a minister who accepts full responsibility for its work. He is subject to examination in parliament, having to explain general policy in debate and to answer questions on even minor details of administration. Over finance, too, there is strict control: the Treasury ascertains that money is spent economically and within the limit authorised by parliament.

But this form of accountability has certain inherent snags. First, parliament is basically a forum for discussing major political issues rather than for dealing with administrative details. In any case, MPs are laymen without the necessary technical knowledge to supervise. Second, parliament would be overworked if it tried to exercise detailed control. Third, frequent questions in parliament on the decisions of civil servants can lead to their taking a 'play-for-safety' attitude or the line of least resistance. The opposition probes mainly with political ends in view; civil servants are therefore hardly likely to follow a bold, imaginative policy which, should it fail, could excite considerable criticism, when by taking an alternative middle-of-the-road line they can settle for a less troubled life. Fourth, Treasury control over finance is restrictive in character. Whereas private enterprise only requires proposed expenditure to be justified by overall profits, the Treasury insists that each individual item of service be provided at the lowest possible cost.

To sum up, accountability clashes with economic efficiency. Thus the government-department form of organisation is most appropriate for dealing with: (i) community and public goods which are basically non-commercial in that, because charges are not levied, their cost is covered by taxation; (ii) in providing merit goods, where only nominal charges are levied; and (iii) where functions are basically limited to supervision or control or both, e.g. foreign affairs and tax collection.

12.5   Local authorities

(a)   The advantages of local-authority provision
While some local authorities may provide trading services, e.g. municipally owned transport, their activities are mostly concerned with community, public and merit goods, e.g. environmental services, roads, parks, education and housing. Making local authorities responsible has certain advantages:

(i) *Those who run local services are local people responsive to local needs and attitudes*
In contrast, administration from Whitehall makes for uniformity throughout the country.

(ii) *It allows close contact between the governed and those who govern*
When administration is at the local level people can get in touch directly with those who make decisions and the officials who carry them out.

(iii) *It facilitates continuity of policy*
Although councils are usually run on party lines, most decisions are of an administrative character, so that only on a few issues, e.g. comprehensive schooling, is there a political division. Councillors can, therefore, work together fairly well to decide the best long-term policies for the community.

(iv) *It provides for a division of power between Whitehall and town hall*
Thus a local protest, e.g. about a proposed airport or motorway, serves to remind the central government that its decisions must respect local feelings.

(v) *Local authorities lighten the burden of administration*
This would otherwise put a severe strain on Westminster and Whitehall.

(b)   Functions of local authorities
For purposes of local government, England and Wales are divided into counties, which in turn are divided into county districts. Below these, *except* in the metropolitan counties and London Boroughs, there are usually parishes (which provide minor local amenities, such as sports facilities and village halls, and may monitor planning applications being considered by the district councils).

   As a general rule the *counties* provide those services which are best administered over a wide area or which are most economically provided on a large scale, e.g. education, police, roads, structure planning and fire services. The *county districts* are basically local planning, refuse collection, sanitation and housing authorities.

   Local government services are provided on the authority of parliament. Most of their functions are *obligatory,* and action can be taken should they be neglected. Others, such as libraries and trading services, are *permissive:* the authority undertakes them at its discretion.

The work of local authorities can be classified under five main headings:

(i) *Protection*
E.g. the police and fire services.

(ii) *Regulation and control*
E.g. the licensing of cinemas, the inspection of weights and measures and of food and drugs, and consumer protection.

(iii) *Personal*
Services for individuals directly, e.g. education, housing and the care of children and old people.

(iv) *Environmental*
The maintenance and improvement of people's surroundings by providing public-health and sanitary services (e.g. refuse disposal), roads, street-lighting, parks, museums, libraries, community centres, etc.

(v) *Trading*
Services provided on a commercial basis, e.g. passenger transport and entertainments.

(c) The local authority as a form of business organisation
Responsibility for carrying out the functions of the particular local authority rests with the *council,* the elected body.

But, apart from meeting infrequently, councils are normally too large to consider business in detail. Nor are any councillors the equivalent of the full-time ministers at Whitehall who take charge of departmental policy. In local government this function is performed by *committees.* Certain committees are required by law to be appointed, e.g. those for finance, social services and education; the rest are decided by the council. Committees are usually given executive as well as advisory functions.

The key figure in a committee is the *chairman*: not only is he in close and regular contact with the corresponding chief officer, but he instigates much of the committee's work; he is the nearest equivalent to the chairman or managing director of a company.

The policies decided upon by the council and its committees are put into effect by paid *permanent officers*. Since the only compulsory appointments are those of education officer and director of social services, councils have great flexibility in planning their management structure. In practice most appoint a chief executive, treasurer, engineer, surveyor, housing manager and parks superintendent, according to the services operated.

(d) Paying for local services
Every year the treasurer of each council submits an estimate of expenditure for the coming year to the finance committee, which considers these estimates and

Where each £ comes from (pence)   How each £ was spent (pence)

	9	Social services
Local taxes (1993 Council tax and Uniform Business Rate) — **35**	8	Police
	2	Fire service
	34	Education
Central government grants — **38**	5	Libraries, museums, parks, environment
	7	Highways & planning
	2	Refuse
	12	Other services
Miscellaneous Fee, interest, Rents trading surpluses etc. — **20**	13	Housing
Deficit — **7**	9	Debit interest

TOTAL
£42,587
million

Figure 12.1   Local authority current revenue and expenditure, 1990–1

decides how the money is to be raised. By far the largest single item of spending is education (Figure 12.1). Other major items are highways, police and social services. Spending is of two kinds, capital expenditure and current expenditure.

Capital expenditure is incurred on such items as new houses, schools, roads, etc. The benefits from these projects extend far into the future, and it would be unfair to throw their full cost onto the ratepayers of one particular year. Such expenditure is therefore covered by long-term loans. Councils borrow mainly on the open market, but where they find this difficult they can obtain funds from the Public Works Loans Board, which is financed by central-government funds.

Current expenditure refers to the everyday spending necessary to run services (e.g. housing administration and repairs, teacher's salaries, student grants) and to pay interest charges. Obviously, if a council is to pay its way it must balance current expenditure with current income from user charges, government grants and the Council Tax.

Charges have the advantage that those who benefit from the service pay at least a part of its cost, as with rents for housing, admission fees to swimming-baths, etc. But while some trading services make a profit (especially lotteries), most local-government activity is concerned with providing community, public and merit goods whose cost has to be covered mainly from taxation, either central or local.

Government grants, which now provide 51% of the revenue of local authorities, are made for the following purposes:

(i)   To offset the defects of the limited yield of the Council Tax.
(ii)   To assist with services of national concern, e.g. roads, police, education.
(iii) To ensure a minimum standard in the provision of such services.
(iv)   To help the poorer authorities.
(v)   To encourage local authorities to provide services above the minimum required.
(vi)   To assist in special emergencies, e.g. floods.

Grants should achieve the above objectives without destroying local control or initiative. But excessive local government expenditure must not prejudice its control of public sector spending as part of its overall economic objectives. It has therefore had to introduce 'rate-capping' to curb over-spending authorities.

The result is that there are three main types of grant:

(i)   *The revenue support grant* to assist with services generally, especially for the poorer authorities.
(ii)   *Specific grants*, some on a percentage basis, e.g. police (50%), highways (50–75% according to the class of road), and others at so much per unit (e.g. housing).
(iii) *The non-domestic rate grant* derived from the Uniform Business Rates levied on business premises according to the annual rent which might reasonably be expected from the property. The proceeds are pooled, and the local authorities' exact allocation is decided by the central government.

The Council Tax

The amount by which the area authorities' spending exceeds revenue from charges and grants has to be covered by levying a Council Tax on households. This replaced the unpopular Community Charge ('Poll Tax') in April 1993. In essence it is a reversion to the old rating system but based on the April 1991 *capital* value of the dwelling instead of a notional annual letting value for which market evidence was deficient. Values are divided into eight bands, A to H, with all households in the same band paying the same amount of tax, but increasing upwards to H. There is a 25% reduction for single occupants.

12.6   The public corporation

The public corporation form of organisation evolved with the nationalisation programme of the Labour government after World War II. The concept envisaged affording these industries freedom to pursue and develop their enterprises according to the commercial objectives associated with the market economy, with some degree of public accountability, but still free from day-to-day questioning in Parliament and Treasury control (as with the government department).

Thus public corporation were made accountable by:

(i)   having a minister responsible for appointing board members and ensuring that the board's policies harmonised with the government's overall economic strategy;

(ii)   allowing parliament to periodically examine how the minister exercised his responsibilities particularly by debating the boards' annual reports;

(iii)   consumers' councils.

But the failure of the nationalised industries to match the enterprise of the market economy led to their 'privatisation' (see below), and today only the Bank of England, a reorganised Post Office and the rumps of British Rail and British Coal remain.

However there are numerous *other quasi-government bodies* ('quangos'), largely independent but still subject to some degree of public control, even if (as with the Church Commissioners) it only means submitting an annual report to Parliament. While some, e.g. the Forestry Commission and the Rural Development Commission, are directly concerned with the allocation of resources, many have arisen because the State has increasingly intervened in economic and social affairs, e.g. the Health and Safety Executive, the Equal Opportunities Commission, the Commission for Racial Equality.

12.7   Privatisation

(a) Nature of privatisation

'Privatisation' implies more than the movement of assets from the public to the private sector. Rather it embraces all the different means by which the disciplines of the free market in the provision of goods and services can be applied to the public sector. Thus this 'pushing back the frontiers of the state' covers:

(i)   the transfer of the nationalised industries to private ownership, e.g. British Telecom, British Gas, British Airways, British Airports Authority;

(ii)   selling other state assets, either completely (e.g. Britoil, Rolls-Royce, motorway service areas) or partially (e.g. woodlands owned by the Forestry Commission, council housing);

(iii)   opening-up State monopolies to outside competition, e.g. relaxing licensing restrictions to allow private bus firms to compete with publicly-owned services;

(iv)   'contracting-out' to the private sector services paid for out of public funds, e.g. refuse collection, street cleaning, hospital ancillary services;

(v)   charging beneficiaries for publicly-provided goods and services, e.g. museums, medical prescriptions, school meals, council housing.

(b) Reasons for privatisation

Although the Labour government initiated a form of privatisation when in 1977 it sold a part of the state's British Petroleum shares in order to be less dependent on borrowing to cover its Public Sector Borrowing Requirement, privatisation is really based on the market economy philosophy of the Conservative party. Thus while Mrs Thatcher's first term of office still concentrated on the PSBR objective, privatisation was extended by returning to the private sector firms which had

been recently acquired (e.g. British Aerospace and Cable and Wireless) and by encouraging contracting-out of services. But during her second term beginning in 1983 privatisation measures were extended and integrated in line with her private enterprise views and supply-side policies (see p. 250). It is in this context, therefore, that the advantages claimed for privatisation have to be analysed.

(i) *Reduced burden on the public purse*
As a one-off, short-term measure the proceeds from state asset sales have helped to cover a worrying PSBR. Indeed, where state industries have had recurrent deficits (with debts eventually having to be written off, e.g. coal, railways, airways), there is a long-term relief to the public purse.

(ii) *Freedom from detailed political control*
Some political control over general policy and the scale of borrowing is necessary to achieve a degree of accountability. But some governments have seen public enterprises as legitimate instruments of macroeconomic policy, e.g. countenancing over-manning to preserve jobs in periods of unemployment, holding prices to combat inflation, and restricting investment spending to reduce the PSBR.

Thus managers of the nationalised industries have felt frustrated at not being able to pursue pricing and long-time investment strategies unencumbered by government interference.

(iii) *Improved efficiency through competition in the market*
Economic efficiency must be considered from both the demand and supply sides. The market indicates consumer preferences (e.g. for ownership of council housing as opposed to renting), while competition promotes efficiency in supplying consumer goods and services.

As monopolies, however, the nationalised industries have tended to take some profits in the form of a 'quiet life' since competition does not force Boards to push the industry to the highest possible efficiency. Even at the lower levels, managers may seek to maximise their own empires and budgets.

(iv) *Greater resistance to trade union power*
Where the state is the ultimate provider of funds, wage demands can more easily be pressed by trade unions and conceded even though not justified by profits. Moreover, especially in the basic industries such as coal, rail and steel, wage rises have been granted on a *national* basis. It is suggested that privatisation undermines the ability of militant public sector trade unions to secure high wages and protection of employment through such subsidies from the taxpayer.

(v) *Creation of a property-owning democracy*
The sale of shares has been so arranged that they have as broad an appeal as possible both as regards price and allocation. Special encouragement has been given to employees to buy shares. Although many purchasers have subsequently sold, a much wider public has been introduced to share ownership.

(c) Difficulties of privatisation

In practice, nationalisation failed. Some industries, such as coal and steel, were over-capitalised, resulting in the 'crowding-out' of investment in the private sector. Moreover, profit targets were regularly missed with the government constantly having to write-off debts. Nor did the energy industries in particular show a high sensitivity to such external considerations as air pollution or the protection of the rural environment. Monopoly powers enabled many to base prices on costs with consequent lack of efficiency.

Even so privatisation is no guarantee that these problems will be eliminated, but steps have been taken to introduce some form of competition wherever possible or to introduce devices to ensure that regard is paid to the 'public interest'.

The major difficulty is that while privatisation eliminates direct government involvement in decision-making and responsibility for particular industries, many, especially the 'natural monopolies' (chiefly public utilities) have retained their monopolist and monopsonist positions. This can result in exploitation of consumers by monopoly pricing and inefficiency through lack of competition.

Consequently, where possible, indirect competition has been fostered. For example, gas still has to compete with coal, oil and electricity, while Mercury has been granted a licence to compete in telecommunications with British Telecom. Similarly, Racal's Vodafone is a major competitor in car telephones with British Telecom's Cellnet. The most striking progress has been in the rapid growth of express coach services with reduced fares after competition with the National Bus Company was allowed.

An alternative arrangement has been to grant independence to firms on a franchise basis for a limited period, e.g. regional television companies. Provision is made to prevent mergers and, in reviewing the franchise, consideration can be given to past conduct as regards quality of service and sensitivity to the wishes of the public as well as to the price tendered. The difficulty with this method is that investment may be inhibited by lack of certainty of long-term future operations.

Where some form of competition is difficult to devise, the responsibility for protecting the public interest may rest with a regulatory body. Thus the Office for Telecommunications (OFTEL) acts as a watchdog for unfair practices, though doubt has been expressed as to whether it has sufficient powers to be really effective. Furthermore, price rises are limited to 7 per cent *less* than the rate of inflation. This ensures that the consumer receives some benefit of technical improvements, but encourages efficiency in that British Telecom is allowed to retain any additional cost savings.

In spite of these difficulties, however, we must recognise the radical nature and achievement of the Thatcher government. Until 1977 the public sector was growing and this seemed to be generally accepted. What the Thatcher government did was to re-examine the proper role of the state in the economy, and other countries, e.g. Italy and France, are now doing the same.

Yet not all government activities can be satisfactorily privatised, e.g. education and medical treatment for the majority of people. For these there must be a continuing process of improving their management and accountability, either by giving them independent control over their allocated finance or by regular efficiency scrutinies.
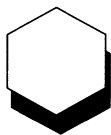
Part VI
Finance and banking

⑬ Finance

13.1 Money

(a) What is money?

It is possible to exchange goods by a direct swap. But barter is rare in advanced economies. Where there is a high degree of specialisation, exchanges must take place quickly and smoothly. Hence we have a 'go-between': money.

Anything which is generally acceptable in purchasing goods or settling debts can be said to be money. It need not consist of coins and notes. Oxen, salt, amber, woodpecker scalps and cotton cloth have at times all been used as money. In fact the precise substance, its size and shape, are largely a matter of convenience and custom. But whatever is used, it should be immediately and unquestioningly accepted in exchange for goods and services. Thus the use of the particular good should be backed by custom, and people must feel that it will retain its value by remaining relatively scarce.

Sometimes an attempt is made to confer acceptability by law. In the UK notes have unlimited *legal tender,* in that a creditor *must* accept them in payment of a debt. But a commodity does not have to be legal tender for it to be money.

A commodity will only be accepted as money if people feel confident it will retain its value.

(b) The functions of money

Money, it is usually stated, performs four functions:

  (i) It is a *medium of exchange* – the oil, as it were, which allows the machinery of modern buying and selling to run smoothly.

 (ii) It is a *measure of value and a unit of account,* making possible the operation of a price system and automatically providing the basis for keeping accounts, calculating profit and loss, costing, etc.

(iii) It is *a standard of deferred payments* – the unit in which, provided its value is stable, loans and future contracts are fixed. Without money there would be no common basis for dealing in debts – the work, for example, of such institutions as insurance companies, building societies, banks and discount houses. By providing a standard for repayment, money makes borrowing and lending much easier.

 (iv) It is a *store of wealth* – the most convenient way of keeping any income which is surplus to immediate requirements. More than that, because money is also the medium of exchange, wealth stored in this form is completely liquid: it can be converted into other goods immediately and without cost. Indeed, it is this 'liquidity' which is the most distinctive

characteristic of money, and it results in money's playing an active rather than a merely neutral part in the operation of the economy.

(c) The supply of money
The supply of money consists of:

(i) *Coins and notes*
Since these are regarded as the small change of the monetary system, sufficient coins and notes are always provided for the everyday convenience of the community.

(ii) *Bank deposits*
While purchases of everyday goods – bus-rides, cigarettes, petrol, etc. – are usually paid for in coins and notes, about 80% (in value) of all transactions are effected by cheque. When a person writes a cheque, the bank is instructed to transfer deposits in his or her account to the person to whom money is owed. Bank deposits therefore act as money. How deposits are 'created' by banks is described in the next chapter.

   The acceptance by the government of the view that the money supply is an important influence in the economy has required that it be defined so that it can be measured and monitored as a guide to policy. Two classifications have been adopted.

   'Narrow money' refers to money balances which are readily available to finance current spending, that is, for transactions purposes. The chosen monetary target is now M_0 which consists of notes and coin held by the public and banks, plus banks' holding of cash (till money) and their operational balances at the Bank of England.

   'Broad money' refers to money held for transactions purposes and money held as a liquid asset. It reflects the private sector's holdings of assets which could be converted with relative ease and without capital loss into spending on goods and services. Here the chosen target is M_4 which consists of M_0 plus all private sterling deposits (sight and time) held in UK banks and building societies.

13.2   The provision of liquid capital

(a) The need for liquid capital
Where expenditure exceeds the receipts of firms or of the government, the deficit has to be bridged by borrowing. Such funds come from the community, which lends savings. Saving represents refraining from spending on consumer goods, thereby setting free resources for the production of capital goods required by firms or for additional expenditure by the government.

(b) Markets for liquid capital
The market is the institution which brings borrowers and lenders together, making funds available to firms and the government at a price – the rate of interest. But, because finance is required by different types of firm and by the government for different purposes and for different periods of time, there is a great variety in

the types of loan available and in the institutions providing or arranging such loans. Nevertheless, markets can be classified into two broad groups: (i) the *money markets* (dealing in short-term loans) and (ii) the *capital market* (where medium-and-long-term capital is raised). The joint-stock banks (the major source of firms' working capital) and the Bank of England (which exercises a general control over the availability of finance) are discussed in Chapter 14.

None of the money markets nor the capital market are formal organisations in that buyers and sellers meet regularly in a particular building to conduct business. Instead they are merely a collections of institutions which are connected, in the case of the money markets by dealings in bills of exchange and short-term loans, and in the case of the capital market more loosely – through channelling medium- and long-term finance to those requiring it. Moreover, as we shall see later, within each market there is a high degree of specialisation.

Because it is such a large borrower, the government's requirements tend to dominate these markets, affecting the rates which have to be paid on short- and long-term loans. The complete structure is shown in simplified form in Figure 13.1.

13.3   Money markets

(a) The discount market
Bills of exchange are an important source of short-term finance – the commercial bill for firms, the Treasury bill for the government. The discount market comprises the institutions linked by dealings in bills – discount houses, merchant banks acting as acceptance houses, commercial banks and the Bank of England.

It is customary in foreign trade for an importer to be allowed a period of grace, usually three months, to pay for goods. This is arranged through a *commercial bill of exchange*.

Suppose *A* in London is exporting cars worth £100,000 to *B* in New York. When he is ready to ship the cars, he draws up a bill of exchange, as shown in Figure 13.2. This is sent to *B*, together with copies of the shipping documents to prove that the cars are on the ship. *B* accepts the bill by writing 'Accepted' and his signature across the face of the bill, and then returns it to *A*. This acceptance of the bill by *B* is necessary before the original *bill of lading,* the documentary title to the cars, is handed over.

A can now do one of three things: (i) hold the bill until it matures; (ii) endorse the bill and then get a merchant to whom he is indebted to take it in settlement; or (iii) sell the bill, usually to a discount house.

(i) *Discount houses*
Probably *A* will choose the latter course. So, after endorsing it, he takes it to one of the *London discount houses*. The exact amount paid for the bill will depend on the length of time to maturity, the prevailing short-term rate of interest and the opinion of the discount house as to B's financial standing. If the bill has three months to run and the prevailing rate of interest on that class of bill is 6%, the discount house will pay about £9,850 for it. This process is known as
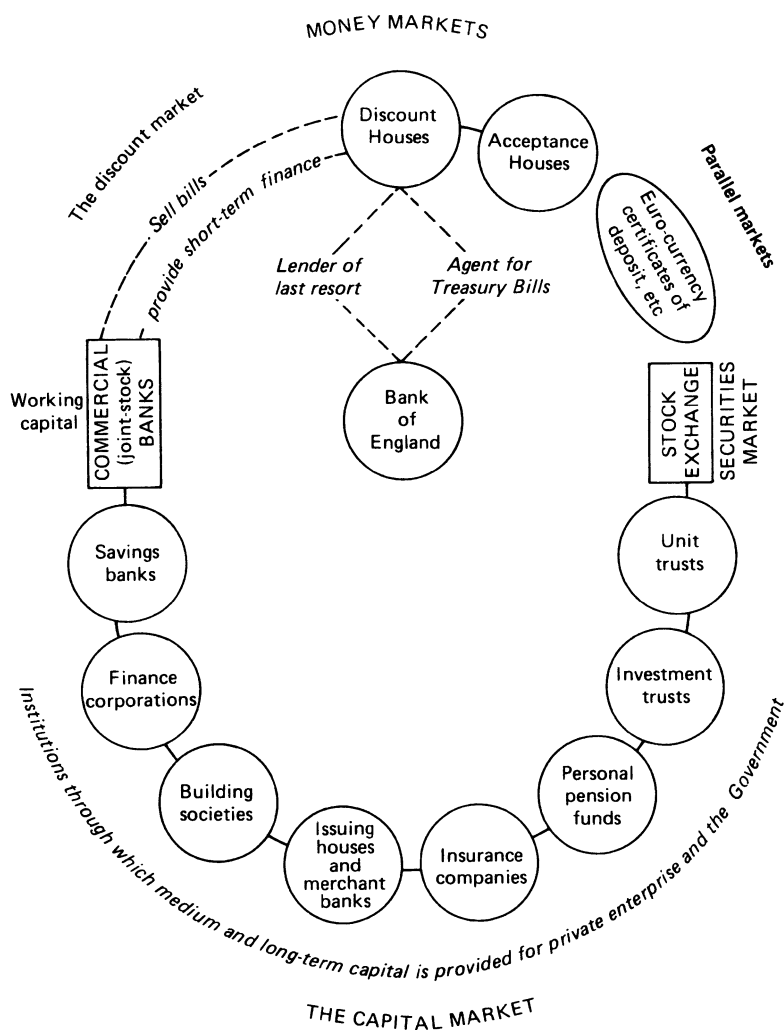
Figure 13.1 The provision of finance in the United Kingdom

'discounting'. Thus, while *A* quickly regains liquidity by selling the bill, *B* obtains three months' credit, during which time he will probably sell the cars.

Bills are not usually held for the full three months. Instead after about a month, they are sold in 'parcels' to the commercial banks, who like to have so many maturing each day.

(ii) *Acceptance houses*

If *B* is a well-known firm of high financial standing, the accepted bill is, from the risk point of view, almost as good as cash. However, as bills are drawn on firms

```
┌─────────────────────────────────────────────────────────────┐
│                                                             │
│   £100,000                                    A's address    │
│                                                             │
│                                               25 April 1994  │
│                                                             │
│              Three months after date, pay to me or to my    │
│              order One Hundred Thousand Pounds, value       │
│              received.                                      │
│                                                             │
│                                                             │
│                                               (Signed) 'A'   │
│                                                             │
│                                                             │
│                                                             │
│      To 'B',                                                │
│           B's address.                                      │
│                                                             │
└─────────────────────────────────────────────────────────────┘
```
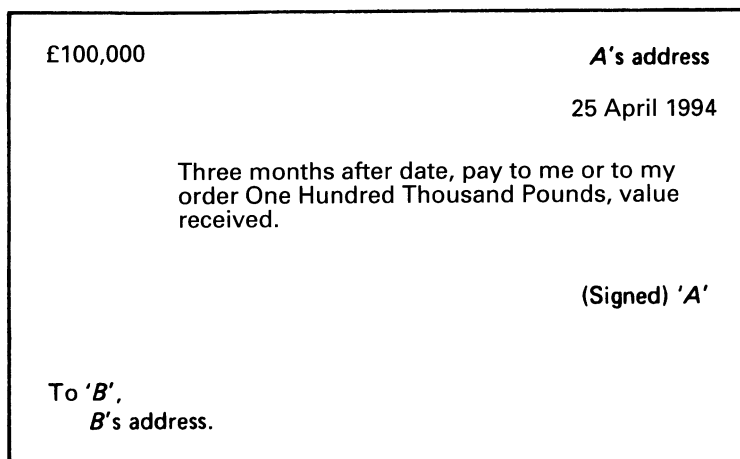
Figure 13.2 A commercial bill of exchange

in all parts of the world, little may be known about B. Thus the discount house is either reluctant to discount the bill or will only do so at a high rate of interest. The difficulty can be overcome by getting a firm of international repute to 'accept' responsibility for payment should B default. It is obvious that any firm accepting such a bill must have adequate knowledge of the creditworthiness of the trader
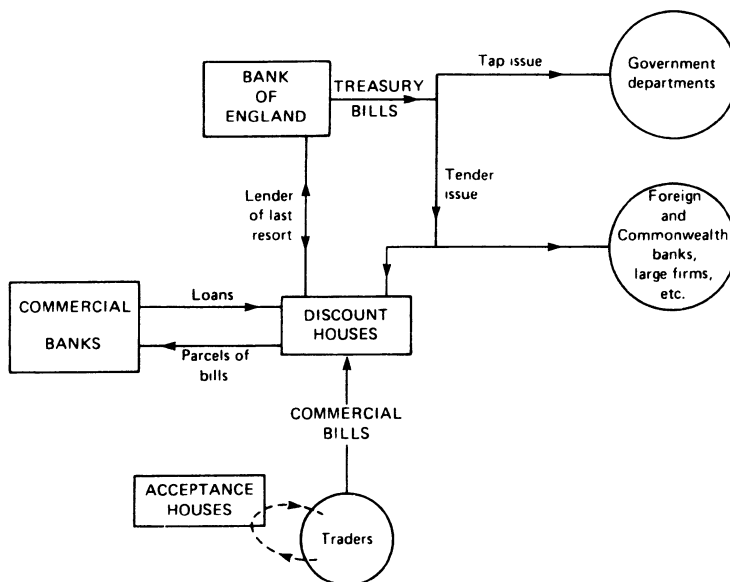
Figure 13.3 Operations of the discount market

upon whom the bill is drawn. Such knowledge is possessed by the merchant banks, such as Lazard, Barings and Rothschild, who commenced as traders but later specialised in financing trade in particular parts of the world. In their capacity of accepting bills such merchant banks are known as *acceptance houses*. For the service, they charge a small commission of about $^3/_4\%$, which is paid willingly because the rate of discount on a 'bank bill', i.e. one bearing the name of an acceptance house, is lower than on a 'trade bill' (a bill accepted only by a trader) or on a 'fine trade bill' (where the merchant is of good standing).

The business of accepting has now declined. Originally this was the result of the diminished use of the trade bill in international trade as the commercial banks competed through the cheaper method of the 'reimbursement credit'. With this, *B*, an importer in New York, asks his own bank to obtain an acceptance credit in London by making itself responsible for payment. Thus the London bank or acceptor there has only to satisfy itself as to the financial standing of the New York bank. This simpler procedure means that reimbursement credits can be granted at very low rates. Today the commercial bill is mainly used as a means of raising finance internally (see below).

The decline in the use of the commercial bill in international trade initially coincided with a large increase in government borrowing through Treasury bills. A *Treasury bill* is really a bill of exchange drawn by the Treasury on itself, usually for a period of three months (ninety-one days), though occasionally two-month bills (sixty-three days) are issued. Treasury bills are only issued in high denominations and so are primarily for institutional investors.

Recent developments have had repercussions on both the discount houses and merchant banks. The restrictions on the lending powers of the banks before 1971 led to the development of other means of short-term borrowing, e.g. internal commercial bills, local authority bills, certificates of deposit, etc. Furthermore, the government has reduced its dependence on short-term borrowing through Treasury bills.

Dealings in these short-term instruments are now, therefore, the mainstay of the *discount houses*, though they still tender for the reduced weekly offering of Treasury bills.

The functions of the *merchant banks* have also changed. The work of accepting is not required for Treasury bills or for most of the new short-term instruments since the standing of the borrower is generally known to be first class. Instead, they arrange and underwrite new issues, advise on the terms of 'take-overs' and mergers, and pay dividends to stockholders as they fall due. They compete in specialist fields, e.g. property development, dealing in securities, domestic banking, and also act as trustees and manage investment portfolios. Other functions have resulted from their overseas trading connections. Thus they have important business in the bullion and foreign exchange markets.

(iii) *The commercial banks*
The commercial banks fulfil two main functions in the discount market – providing the discount houses with funds and holding bills to maturity.

The discount houses do not themselves have sufficient finance to buy all the bills, commercial and Treasury, offered them. They overcome this difficulty, however, by borrowing money from the commercial banks at a comparatively low rate of interest. Then, by discounting at a slightly higher rate, they make a small profit. The banks are willing to lend at a low rate because the loans are of short duration, often for only a day, and need not be renewed if there is a heavy demand for cash from their ordinary customers. For the discount houses, the trouble involved in the daily renewal of this money at call and the slight risk of its non-renewal are compensated for by the comparatively low rate of interest charged.

The commercial banks can earn a higher rate of interest by themselves holding bills for a part of their currency. However, except on rare occasions, they do not bid for them directly but buy them from the discount houses when they still have about two months to run.

(iv) *The Bank of England*

The Bank of England enters the discount market as follows.

First, it is the agency by which the government issues Treasury bills. This issue is achieved by two methods, 'tap' and 'tender'. Government departments, the National Savings Bank, the Exchange Equalisation Account, the National Insurance Fund and the Bank of England Issue Department, all of which have funds to invest for a short period, can buy what bills they want at a fixed price, i.e. 'on tap'. This price is not published.

The discount houses and other purchasers (such as Commonwealth and foreign banks) can obtain their issue by 'tender'. Every Friday, the Treasury, acting through the Bank of England, invites tenders for a specified amount of bills, usually £100 million each week.

Second, the Bank of England is the 'lender of last resort'. When the discount houses are pressed for money because the commercial banks will not renew their 'call money', the Bank of England will make finance available to them (see p. 199).

(b) Parallel money markets

As a result of restrictions placed on bank lending, new secondary markets in short-term loans developed to meet the specific requirements of particular borrowers and lenders. Indeed the existence of such markets has encouraged funds to be lent short-term, since they have made it easier for lenders to regain liquidity.

The following are the most important of these comparatively new markets:

(i) *Sterling interbank market*

This is a market bringing together all banks, including merchant banks, British overseas banks and foreign banks, so that those having funds surplus to their immediate requirements can lend to those having outlets for short-term loans or requiring greater liquidity. It is described as a 'wholesale' market, as opposed to a

'retail' market where funds are collected directly from the public, e.g. by building societies, clearing banks, National Giro.

The going London Inter-Bank Offered Rate (LIBOR) is the key rate for other short-term loans.

(ii) *Local authority market*

Local authorities borrow on the open market. Short-term loans bridge the time difference between expenditure and revenue. Brokers now exist for placing with them short-term funds of banks, industrial and commercial companies, charitable funds, etc. Such brokers also deal in longer-term local authority bonds.

Today the market is integrated very closely with the interbank market, as funds from the latter are very often deposited with local authorities.

(iii) *Negotiable certificates of deposit market*

Certificates of deposit enable the banks and building societies to obtain 'wholesale deposits' for periods from three months to five years. They are like bills of exchange but drawn on themselves. Since they are for a longer period than an ordinary time deposit, they facilitate medium-term lending. For the lender they offer a higher rate of interest, while the market in them means that they can be sold whenever cash is required.

(iv) *Eurocurrency market*

Eurocurrency deposits are simply funds which are deposited with banks outside the country of origin but which continue to be designated in terms of the original currency. The most important Eurocurrency is the dollar. As a result of the USA's continuing adverse balance of payments, branches of European banks have built up dollar balances as customers were paid for exports. These balances are offered to brokers in London (where interest rates have been higher than in New York), and are placed mainly with companies or banks (e.g. Japanese) operating on an international scale to finance foreign trade or investment. While the dollar still dominates the market, other European currencies are now dealt in, chiefly the Deutschemark, the Swiss franc and the Japanese yen. The availability of such funds has complicated credit control by the Bank of England (see p. 199).

(v) *Other markets*

Smaller specialist markets have developed in *finance-house deposits* and *intercompany deposits*. Thus finance houses have obtained funds by issuing bills which are accepted by banks and discount houses. Similarly, in periods of tight credit, firms which are short of finance turn to other companies which temporarily have funds to spare.

13.4   The capital market

Whereas the money markets developed to supply short-term finance to trade and the government, industry obtains most of its 'working' capital from the commer-
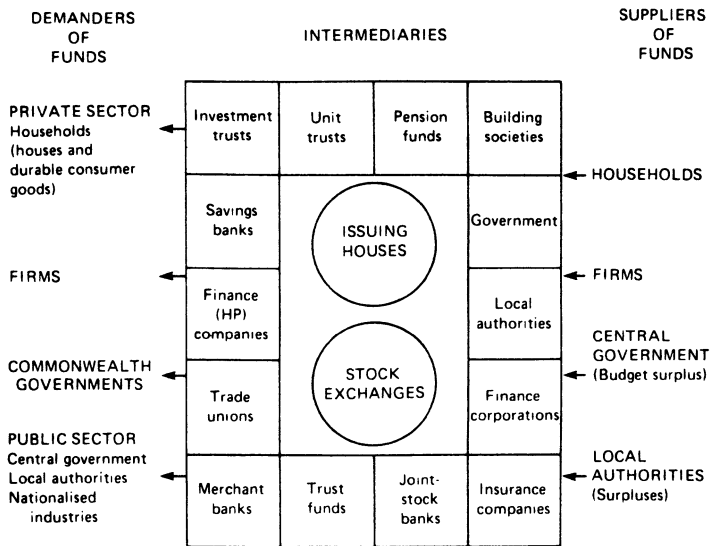
Figure 13.4   The capital market

cial banks (see Chapter 14). But long-term capital for both the public and private sectors is obtained through the capital market.

As can be seen from Figure 13.4, this consists of, on the one hand, the suppliers of long-term capital and, on the other, those requiring such capital, the two being connected by a number of intermediaries, usually of a specialist nature. Some of these intermediaries have already been described; here we look briefly at the others.

(a) Insurance companies
Insurance companies receive premiums for insuring against various risks. Some of these premiums, such as those for insuring ships and property, are held only for relatively short periods – having, apart from the profit made, to be paid out against claims. But with life insurance, endowments and annuities, premiums are usually held for a long time before payments are made. Hence insurance companies have large sums of money to invest in long-term securities. These investments are spread over government and other public stocks, the shares and debentures of companies, property and mortgages. Today 'institutional investors', of which insurance companies are the most important, supply the bulk of savings required for new issues.

(b) The Girobank
The Girobank carries out all essential banking services, mostly for private customers. Its strength is that it operates through the countrywide network of Post Offices. All records are kept at the computerised centre in Bootle, Lancashire. It is now part of the Alliance and Leicester Building Society.

(c) Foreign banks
Branches of over 400 foreign banks are now located in London. Early on they thrived because they were free of the strict credit controls imposed by the government on the clearing banks. Their more recent expansion reflects the development of the international banking system. While they carry out normal banking functions, their work is heavily concentrated on foreign exchange dealing.

(d) Trust, pension and trade-union funds
All these accumulate income which is re-invested in government securities, shares, property, etc.

(e) Building societies
The main functions of building societies are still the collection of retail deposits from the general public and the granting of long-term loans for the purchase of dwellings for owner-occupation. In recent years they have supplied cheque-books, credit cards and other services to depositors, thereby competing with the banks.
   The Building Society Act 1987 allowed them to convert into companies (e.g. Abbey National), own property (mostly residential), grant second mortgages and unsecured loans up to 10 per cent of their total lending and to provide a variety of financial services connected with house purchase, e.g. arrange surveys, insurance.

(f) Finance corporations
Apart from 3i (see p. 65) owned by the clearing banks and the Bank of England, there are a number of similar consortiums, e.g. the Agricultural Mortgage Corporation, the Export Credits Guarantee Department of the Department of Trade and Industry, and the British Screen Finance Consortium which provide finance in their specialist fields.

(g) Finance houses
These were originally independent companies set up to borrow from the public and banks in order to finance hire purchase of both consumer goods and machinery. Today the industry is dominated by the larger commercial banks. For instance, Mercantile Credit is now part of Barclays Bank, and the United Dominion Trust a part of the Trustee Savings Bank.

13.5   Markets in securities: the Stock Exchange

(a) History
By the second half of the seventeenth century there was a recognisable market for dealing in securities. This was gradually formalised and in 1773 the Stock

Exchange occupied its first settled premises and from 1803 published its *Official List* of prices. From 1908 its organisation was based on a separation between 'brokers' and 'jobbers'. Brokers acted on behalf of their clients buying from and selling shares to jobbers, the dealers in the shares. This 'single capacity' requirement was designed to protect clients. Whereas brokers worked on a commission basis, jobbers relied on profits from their dealings.

(b) Recent influences
This cosy arrangement was jolted in the early 1980s by two developments. First, in 1970 government policy put greater emphasis on extending competition throughout the economy, and the Office of Fair Trading frowned upon fixed commissions as a monopolistic practice. Furthermore, the government was keen to maintain and even develop London's invisible earnings capacity (see p. 293). The abolition of fixed commissions in New York in 1979 made dealing costs for British institutions lower there than in London, while the ending of exchange control in the same year meant that British investors were unhampered in investing in foreign securities.

The second development was technological – the introduction of electronic information and communication systems. This meant that changes in security prices in a dealing centre in one part of the world could be transmitted and indicated visually on screens in other centres. Thus the three leading centres, Tokyo, London and New York, became one market in which, because of the time difference, dealing took place over almost the twenty-four hours of the day.

Thus the pressure was on the Stock Exchange to revise its fixed commission arrangements and to adopt a less parochial outlook. The actual changes took place on 27 October 1986 and produced such an immediate upheaval that it was referred to as 'Big Bang'.

(c) 'Big Bang'
The Stock Exchange agreed to end fixed commissions. But since this would have forced many brokers out of business it was necessary to end the 'single capacity' rule and allow members to act in a dual capacity as agents for both clients and dealers. The main dealers are termed 'market makers'.

Market-makers negotiate their own commissions for buying and selling shares, and on the larger orders put through by the institutions can offer attractive terms. While private investors can also negotiate terms, the size of their business is not deemed profitable by the market-makers and commission rates remain much as they were before 'Big Bang'. This leaves room for brokers to earn a respectable living by acting as the retailer for the private investor, providing a personal contact and offering advice and even research.

(d) Further developments
'Big Bang' proved to be the catalyst for even more far-reaching developments. The government's desire to establish London as an international trading centre necessitated making dealing costs more competitive and so stamp duty on shares was reduced to $\frac{1}{2}$ per cent. But to trade in competition with the larger Japanese

and American firms UK dealers had to have access to considerably more capital in order to carry stocks of securities. Thus firms had to merge or, more usually, were taken over by larger financial institutions, such as the merchant banks.

In this, however, 'Big Bang' simply gave impetus to the movement which was already taking place of linking related services in one firm. We can take the major clearing-banks as an example. As mentioned above, they have already acquired an interest in hire purchase finance and have now followed up their provision of loans for house purchase by acquiring firms of estate agents. The idea has been extended into stock-broking. Market-making, however, has been left to the merchant banks who are less interested in the retail side of finance.

We still refer to the 'Stock Exchange' although dealing is no longer on the 'floor of the House'. Instead there is the Stock Exchange Automated Quotation system (SEAQ) which is the electronic market-place of the London stock market. Information from the sixty-four market-makers on prices and deals made is fed into SEAQ and displayed on screens. This enables the market-maker to quote a selling price and a lower buying price. The difference will be larger when the shares are only dealt in infrequently or where the sale of comparatively few shares can lead to a large fall in price. Unfavourable news, such as a poor monthly balance of payments figure, will cause him to lower prices as a precautionary measure. These new prices would be recorded on the SEAQ screen for the rest of the market.

The government has a strong interest in the integrity of the market but opted for allowing the City to regulate itself rather than impose centralised control. The Secretary of State for Trade and Industry appointed a Securities and Investments Board (SIB) which oversees six Self-Regulating Organisations (SRO) and City dealers have to belong to one of these. The Stock Exchange, as a SRO, issues guidelines to members and ensures that these rules are adhered to.

(e) Economic functions
As the two main UK markets in securities are still subject to the discipline and regulations of the Stock Exchange Council, both are included under the omnibus term of 'the Stock Exchange'.

Critics of the Stock Exchange tend to ignore its real functions and to concentrate on its speculative aspects. It is true that the facilities offered do provide openings for speculation. The fortnightly account allows a speculator to buy securities at the beginning and sell within fourteen days without ever having to put up any money. A speculator who buys securities because he thinks the price will rise is said to be a 'bull'; he hopes to sell them at a profit before the end of the account. On the other hand, a 'bear' sells securities he does not possess because he expects the price to fall before they have to be delivered.

The difficulty concerning speculation is that both optimism and pessimism are contagious and so the market becomes extremely susceptible to both over-confidence and panic. Indeed, expectations are 'self-fulfilling': people who expect the price of securities to rise bid for them, thereby sending up their price, and vice versa. The result is that the prices of stocks and shares may be written up or down not through changes in their earnings prospects but simply through waves of confidence or mistrust.

Even so, we must not forget that some speculation may be advantageous. Expert professional operators tend to steady prices through their function of holding stocks (see below). This also permits securities to be bought and sold at any time, thereby making them more liquid. The great difficulty occurs in distinguishing harmful speculation from genuine investment, for with all investment there is a certain element of risk. In any case the magnitude of the speculative business must not be overestimated. Most purchases represent genuine investment conducted on behalf of investment trusts, insurance companies, pension funds, building societies and private individuals.

The truth is that, for the following reasons, an organised market in securities is an indispensable part of the mechanism of a market economy.

(i) *It facilitates borrowing by the government and industry*
If people are to be encouraged to lend to industry and the government by the purchase of securities, they must be satisfied that they will subsequently be able to sell easily those investments which they no longer wish to hold. Such an assurance is afforded to any holder of a fairly well-known security by the Stock Exchange, for it provides a permanent market bringing together sellers and buyers.

Thus, indirectly, the Stock Exchange encourages savers to lend to the government or to invest in industry. Indeed, if a new issue receives a Stock Exchange quotation, the chances of its success are considerably enhanced.

(ii) *Through the market-makers, it helps to even out short-run price fluctuations in securities*
By holding stocks of shares, a dealer provides in the short run a buffer against speculation by outsiders. This is because he does not merely 'match' a buyer with a seller but acts like a wholesaler, holding stocks of securities. Since he usually specialises in dealing in certain securities, he obtains an intimate knowledge of them. Thus when the public is pessimistic and selling, he may be more optimistic in his outlook and consider that the drop in price is not likely to continue. He therefore takes these securities on to his book. Similarly, when the public is rushing to buy he will, when he considers the price has reached its zenith, sell from his stocks. The effect in both cases is to even out the fluctuations in price, for, in the first case, he increases his demand as supply increases, and, in the second, he increases supply as demand increases.

(iii) *It advertises security prices*
The publication of current Stock Exchange prices enables the public to follow the fortunes of their investment and to channel their savings into profitable enterprises.

(iv) *It protects the public against malpractices and fraud*
With dealers acting in a dual capacity, the previous safeguard of a client that his broker acted solely on his behalf was lost. Under the new arrangements there are two safeguards, the open display of prices on the SEAQ screen and the

regulations of the Stock Exchange Council as a SRO. The Council insists on a high standard of professional conduct from its members. Should any authorised member default, the investor is indemnified out of the Securities Association compensation fund.

The *Official List* of securities indicates that the Stock Exchange considers shares are reputable. Permission to deal is withdrawn if any doubts arise about the conduct of a company's affairs.

(v) *It provides a mechanism for the raising of capital by the issue of securities*
While the Stock Exchange is essentially a market for dealing in 'old' securities, the success of a new issue to raise capital is enhanced if a promise can be made of a Stock Exchange quotation for it. More directly, brokers and dealers will actively arrange for certain clients to provide capital for firms wishing to expand (see p. 67).

(vi) *It reflects the country's economic prospects*
The movement of the market acts as a barometer which points to the economic prospects of the country – whether as 'set fair', or otherwise!

⬡14 Banking

14.1   The British banking system

(a) Types of banks in the UK

(i) *The commercial or joint-stock banks*
These banks are dominated by the 'Big Five': Lloyds, Barclays, National West-minster, Midland (now part of the Hong Kong and Shanghai Banking Corpor-ation) and the Trustee Savings Bank. They operate through a network of branches throughout the country. Their importance in the financial system stems from the fact that most of their business is conducted by way of cheques which, through their central clearing arrangements, enables them to economise in the use of cash and so 'create' credit.

(ii) *The central bank*
This is the Bank of England, which, on behalf of the government, exercises ultimate control over the financial system.

(iii) *Merchant banks* (see p. 182).

(iv) *Foreign and commonwealth banks having branches in the UK*

(v) *The National Savings Bank*

(vi) *The Girobank* (see p. 186).
Easily the most important of the above are the joint-stock banks. But attention must also be given to the Bank of England, especially with regard to its influence on the operation of the joint-stock banks.

14.2   Joint-stock banks

(a) The cheque system
Banks are companies which exist to make profit for their shareholders. They do this by borrowing money from 'depositors' and re-lending it at a higher rate of interest to other people. Borrowers are private persons, companies, public corporations, the money market and the government. Other things being equal, the more a bank can lend, the greater will be its profits.

People who hold a current account at a bank can settle their debts by cheque or credit card, a very convenient form of payment. But the use of cheque is advantageous to banks. Thus, to advertise their business, induce customers to pay by cheque rather than by cash and encourage people to deposit money with them, banks perform many services outside their main business of borrowing and lending money – keeping accounts, making standing-order payments, providing night-safe and cash dispenser facilities, paying bills by credit transfers, purchasing securities, transacting foreign work, storing valuables, acting as executors, etc.

(b) The cheque as a substitute for cash

Cheques lead to a reduction in the use of cash. Suppose that I have paid £1,000 cash into my banking account. Imagine, too, that my builder banks at the same branch and I owe him £500. I simply write him a cheque for that amount, and he pays this into the bank. To complete the transaction, my account is debited by £500, and his account is credited by that amount. What it is important to observe, however, is that in the settlement of the debt no actual *cash* changes hands. A mere book entry in both accounts has completed the transaction.

Perhaps my builder will, towards the end of the week, withdraw some cash to pay his workers' wages. But it is likely that most of his payments, e.g. for building materials, petrol and lorry servicing, will be by cheque. Similarly, while I may withdraw some cash from the £500 still standing to my account to cover everyday expenses, the probability is that many of my bills, e.g. my club subscription, half-yearly rates and mortgage repayments, will be settled by cheque or by transfer directly from my account. Furthermore, even where cash is withdrawn by one customer, this is often compensated for by cash being paid in by others.

With the development of the cheque system the proportion of cash required for transactions has decreased. Let us assume a simple model in which the banks operate free from government control but have discovered that in practice only 10% of their total deposits need be retained in cash to cover all demands for cash withdrawals. In short, only £100 of my original deposit of £1,000 is needed to form an adequate cash reserve.

(c) The creation of credit

It is obvious, therefore, that £900 could be lent by the bank to a third party without my being the wiser. What is not quite so obvious is that the bank can go further than this – and does!

Let us assume that there is only the one bank and that all lending is in the form of advances (see p. 195). When a person is granted a loan by his bank manager, all that happens is that the borrower's account is credited with the amount of the loan, or alternatively that he is authorised to overdraw his account up to the stipulated limit. In other words, a deposit is created by the bank in the name of the borrower.

When the loan is spent, the borrower will probably pay by cheque. If this happens, there is no immediate demand for cash. There is no reason, therefore, why

the whole of my cash deposit of £1,000 should not act as the safe cash reserve for deposits of a much larger sum created by the bank's lending activities.

But the bank must not overdo this credit creation. Our model has assumed that, to be safe, cash must always form one-tenth of total deposits. This means that the bank can grant a loan of up to £9,000. Because it is the only bank, there is no need to fear that cheques drawn on it will be paid into another bank and eventually presented for cash.

The process of credit-creation is illustrated in Figure 14.1. *X* pays £1,000 in cash into the bank. This allows the bank to make a loan of £9,000 to *B*, who now settles his debts of £4,000 and £5,000 by sending cheques to *C* and *D*.
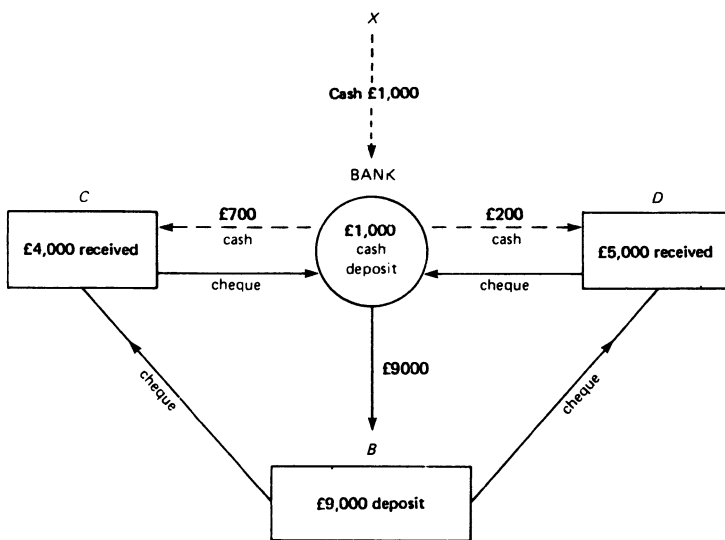
Figure 14.1 How a bank creates credit

These cheques are paid into the bank. *C* withdraws cash rather heavily, in the sum of £700, but this is compensated for by *D*, who only withdraws £200 in cash. This leaves £100 cash – enough to cover the average withdrawal which *X* is likely to make. While these cash withdrawals are being made, other cash is being paid in, thereby maintaining the 10% ratio.

In practice there are many banks, but for the purpose of credit creation they are virtually one bank because they are able to eliminate a large demand for cash from each other by their central clearing arrangements. Moreover, banks keep in line as regards their credit creation. Were one bank to adopt, say, a 6% cash ratio, it would find that, because its customers were making such a large volume of payments to people who banked elsewhere, it would be continually called upon to settle a debit with other banks in cash at the end of the day's clearing, so that its cash reserve would fall below the safe level.

(d) The effect of lending on the bank's balance-sheet

Suppose that the receipt of the £1,000 cash and the loan to *B* are so far the sole activities of the bank. Ignoring shareholders' capital, its balance-sheet will read:

Liabilities	£	Assets	£
Deposits:			
Deposit account	1,000	Cash in till	1,000
Current account	9,000	Advances	9,000
	10,000		10,000

The advance to *B* is an asset; it is an outstanding debt. On the other hand his current account has been credited with a deposit of £9,000 – just as though he had paid it in. It can be seen, therefore, that *every loan creates a deposit*.

X would pay the £1000 into his current account if he himself had an overdraft. But if he had a sufficient credit balance there, he could pay it into a deposit account where it would earn some interest. Such deposits are referred to as 'time deposits' since technically they are subject to a withdrawal notice (though banks usually waive this subject to loss of interest). The loan to *B* would be credited to *B*'s current account and, since it is available for immediate spending, is known as a 'sight deposit'.

(e) The bank's objectives

In practice the structure of the bank's assets is more varied than this. Creating deposits in order to lend at a profit entails risks. First, the loan may not be repaid. Second, and more important, there may be a run on the bank for cash: the original depositor may wish to withdraw his £1,000, or *B*, *C* and *D* may require between them an abnormally large amount of cash. Any suggestion that the bank could not meet these demands would undermine confidence.

Hence, although a permanent cash reserve is essential, a bank must have other lines of defence so that in an emergency it can raise cash easily and quickly. Instead of lending entirely by advances, which a borrower usually requires for at least six months, some loans must be restricted to a shorter period, even to just a day. But, the shorter the period of the loan, the lower will be the rate of interest the bank earns. On the other hand, it wants profit for shareholders to be as high as possible. Thus it is limited in its lending policy both quantitatively and qualitatively. Not only must credit be restricted to a multiple of the liquid reserves but it must also afford adequate *security, liquidity, and profitability.*

As regards *security*, the bank endeavours not to lend if there is any risk of inability to repay. Collateral, e.g. an insurance policy, the deeds of a house or share certificates, is regarded more as a weapon to strengthen its demand for repayment than as a safeguard against default.

Liquidity and *profitability* pull in opposite directions – the shorter the period of the loan, the greater the bank's liquidity but the less it will earn by way of interest. The difficulty is resolved by a compromise: (i) loans are divided among different types of borrower and for different periods of time; and (ii) the different
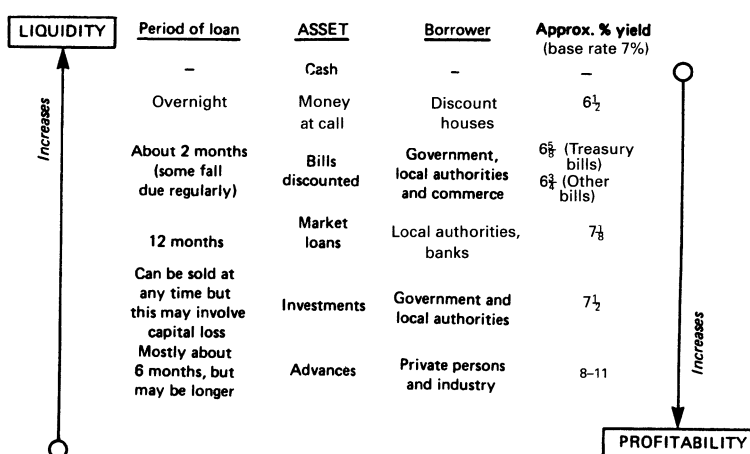
Figure 14.2 The nature and distribution of a bank's main assets, 1993

types of loan are kept fairly close to carefully chosen proportions. In short, the bank maintains a 'portfolio' of assets.

(f) The distribution of a bank's assets

How in practice a bank reconciles the aim of liquidity and profitability can be seen by studying its sterling assets. This is possible because, apart from its cash, buildings and goodwill, loans represent its sole assets. Just as 'sundry debtors' appears on the asset side of a firm's balance sheet, so debt outstanding to a bank represent assets to it. The position is shown in Figures 14.2 and 14.3.

Cash covers: (a) till money, to meet customers' demand for coin and notes; (b) the working balance at the Bank of England to cover any liability on the day's clearing; (c) the non-operational balance of 0.45 per cent of total liabilities which has to be deposited with the Bank of England.

Bills, which are Treasury bills, local authority and trade bills, are obtained chiefly from the discount houses (though some may be discounted directly for customers) and are held for the remainder of their currency – usually two months.

Market loans consist mainly of: (a) money at call and short notice which enables the discount houses to discount bills and hold them for a month or so before passing them on to the banks (see p. 183); (b) loans of less than a year to local authorities; (c) certificates of deposit (see p. 184); (d) short-term loans to other monetary authorities.

Investments are medium- and long-term government securities bought on the open market.

Advances, to nationalised industries, companies and personal borrowers, are the most profitable (1 to 4 per cent above base rate) but also the least liquid of all the bank's assets. The main object of advances it to provide the working capital for industry and commerce. The type of loan preferred is 'self-liquidating' within a period of about six months. A good example is a loan made to a farmer, who

borrows to buy seed and fertiliser and to pay wages, and repays the loan when the harvest is sold. Similarly, a manufacturer may borrow to employ additional labour and raw materials just prior to Christmas in order to increase production. When payment is received for those goods he can repay the overdraft. Borrowers are often allowed to 'roll over' their overdrafts.

Banks also make a large number of 'personal loans', usually modest sums to cover exceptional items of personal expenditure. Repayments are spread over the term of the loan, though interest is charged at an agreed rate on the full amount of the loan for the whole of the period.

At one time banks refrained from providing long-term capital leaving this to the capital market. In recent years, however, they have competed in financing such long-term projects as the construction of offices and the purchase of farms and owner-occupied houses. These fixed assets are the security required.

It must be emphasised that, apart from cash and bank buildings, these assets are covered only by credit created by the bank. For example Treasury bills and government securities are paid for by cheques which will increase the accounts of the sellers. If they are new issues, there is an addition to the government account at the Bank of England; if they are old issues, the bank is virtually taking over from somebody else a loan already made to the government. In writing these cheques the bank increases its liabilities, for only book-entry deposits cover them. This 'pyramid of credit', created to buy earning assets and to make loans upon a liquid-assets basis, is shown in Figure 14.3.
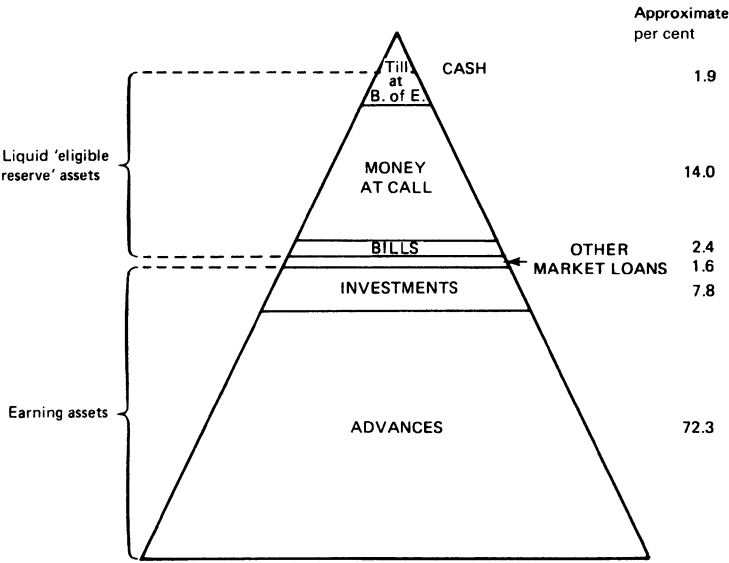
Figure 14.3 The pyramid of bank credit

(g) Modification of the cash-ratio approach

Our explanation of credit creation has followed traditional lines: credit bears a fixed relationship to cash reserves. Today, however, the monetary authorities (i.e. the Bank of England acting as agent for the Treasury) now regard cash simply as the small change of the monetary system, to be varied according to the needs of trade. In any case banks are more concerned with their general liquidity position than with the one item of cash. This originally stemmed from the introduction of the Treasury bill, which through government support became almost as good as cash. Improved markets for loans, e.g. the money markets described in Chapter 13, also increase liquidity to the extent that such loans can be regarded as 'near money'.

It follows, therefore, that if the authorities wish to control the money supply, their attention must be directed to the total assets which the banks hold.

(h) The effects of recent increased competition

The above describes the fundamental credit-creation principle of banking. But the role of the clearing banks in the financial structure of the UK has been revolutionised since 1960 as they have had to respond to increased competition. This has come in four stages:

(1) Overseas banks (now over 400) came to London in the 1960s to share the Eurocurrency market there. In addition to dominating foreign currency lending, they now have some 30 per cent of lending in sterling to British companies and financial institutions.

(2) The Competition and Credit Control policy of 1971 removed restrictions on lending, thereby ending the clearing banks' 'interest rate' cartel. No longer could they simply rely on recycling personal deposits into loans to business. Now each bank had to compete for personal customers, e.g. by extending banking hours, increasing lending (especially through mortgages) and offering higher interest on deposits.

(3) The ending of exchange controls in 1979 removed the remaining restrictions on foreign lending institutions.

(4) The development of information technology has made possible credit cards, point of sales transfers and dealing in securities on an international scale.

Increased competition has meant that bank margins between borrowing and lending rates have been squeezed and many personal customers' services which were previously given free are now charged for. In their turn the banks have had to improve customer relations.

In many of their new ventures, e.g. estate agency, personal mortgage lending, loans to Third World countries and property development finance, the banks' entry into the more competitive market has incurred them in heavy losses.

14.3   The Bank of England

(a) Functions of the Bank of England

The Bank of England, established in 1694, remained a joint-stock company until it was nationalised in 1946. However, nationalisation merely formalised its

position as a 'central bank' – the institution which, on behalf of the government, exercises the ultimate control over the policies of banks and other financial institutions.

The Bank of England performs a number of functions. It issues notes, acts as banker to the government, holds the cash reserves of the commercial banks and exercises a general supervision over them, manages the Exchange Equalisation Account in order to stabilise the sterling exchange rate or to protect the gold and foreign-currency reserves, and co-operates in harmonising monetary policies with those of other central banks and international financial institutions, such as the International Monetary Fund (IMF), the Bank for International Settlements and the International Bank for Reconstruction and Development (the World Bank). Above all, it implements government monetary policy.

(b) Principles of monetary control

The Bank of England, in its role of a central bank, is responsible for seeing that the monetary system of the country is working in harmony with the government's economic policy.

In broad terms, monetary policy means varying the cost and availability of credit. It can be approached in two main ways: (a) by controlling the *supply* of credit quantitatively – that is, restraining lending by financial institutions; (b) by controlling the *demand* for credit by manipulating its price – that is, varying the rate of interest.

(c) Quantitative control through the old system of Competition and Credit Control (CCC)

The system of CCC initiated in 1971 put the emphasis on controlling the supply of credit *quantitatively*. There were four main weapons:

(i) *The minimum reserve assets ratio*

The foundation of the policy rested on the ability of the Bank of England to dictate to the banks and other lending institutions their minimum liquidity ratio.

Each bank was originally required to observe a minimum reserve ratio of $12\frac{1}{2}$ per cent of 'eligible reserve assets' (liquid assets, as defined by the Bank of England) to 'eligible liabilities' (broadly net bank deposits).

(ii) *Open-market operations*

By buying or selling government securities on the open market, the Bank of England can vary the ability of the banks to create credit. Suppose it sells long-term securities. The increase in the supply offered lowers their price (that is, raises the rate of interest) until the total offering has been bought by the banks or their customers. Since cash is necessary to pay for them, the banks' cash balances at the Bank of England will fall. In other words, the liquid reserve assets held by the banks are reduced and, if previously banks had made loans to the maximum possible extent, they will be forced to squeezed their advances.

Alternatively the Bank of England may put the pressure on the short end of the market by varying the weekly offer of Treasury bills. In as much as these bills are

bought initially outside the banks, the cash balances of the bank's customers and hence the cash held by the banks both fall.

(iii) *Special deposits*
The supply of credit can be influenced by calls for special deposits, requiring banks to deposit with the Bank of England a prescribed percentage of their total eligible liabilities. Such special deposits reduce the bank's liquid assets and thus the bank's ability to make loans.

(iv) *Minimum lending rate*
The minimum lending rate (MLR) was the rate of interest at which the Bank of England would help the discount houses as lender of last resort. Although the rate was announced each Thursday, it was only changed for compelling reasons and so its effect was to give some stability to interest rates in general.

At various times the above weapons were supplemented by *requests* to the banks to restrict or discriminate in their lending, *funding* operations to reduce overall liquidity by converting government short-term borrowing into long-term debt, and *supplementary deposits* (the 'corset') which required a bank whose interest-bearing deposits had risen faster than a specified rate to place with the Bank of England a given proportion of the excess.

(d) **Rate of interest control: the present system**
The main purpose of quantitative controls was to reduce the need to raise interest rates, at least in the short term, by causing banks to ration their lending. In practice, quantitative controls only postponed the rise in interest rates needed. At the same time they not only tended to reduce competition but caused funds to move out of the controlled banking sector to uncontrolled institutions, such as the merchant banks, finance houses and foreign banks. Furthermore, the creation of money substitutes (e.g. the commercial bill of exchange) weakened the role of sterling M_3 as a monitor of the money supply. Above all, quantitative controls became inoperable in 1979 when exchange controls were abolished for funds could now be obtained from overseas sources.

Instead of following a monetary base policy, the Bank now adjusts the supply of credit by *influencing* the short-term rate of interest. This, as *lender of last resort,* it can do by making the market short of cash, as follows.

Each day the Bank of England estimates the cash position of the banking system for this knowledge enables it to use open market operations to keep the banks on a tight rein. This is possible because the banks' margins are now so competitive that they have to be fully lent by keeping their non-earning operational cash balances to a minimum. Hence if depositors withdraw an abnormal amount of cash, the banks have to call in their overnight loans (money at call) from the discount houses.

When pressed for cash by the banks, the discount houses turn to the Bank of England. As lender of last resort, the Bank provides finance by buying Treasury bills or eligible bank bills from them. But it does not quote its terms for that would loosen its grip on the reins. Instead the discount houses have to suggest a

price. If the Bank does not like this price, they are forced to lower the price, thereby giving an upward nudge to the seven-day interest rate. Usually, however, the Bank's objective is simply to even out normal fluctuations in the demand for and supply of cash in order to stabilise the rate of interest prescribed by the Treasury.

Twice a day the Bank announces the rates at which it has bought and sold bills. By doing so it indicates to the market what is the appropriate short-term rate of interest.

Should, however, the Chancellor of the Exchequer wish to alter the rate, the Bank of England will signify this in its terms. Usually the Bank operates on the seven-day rate. The commercial banks' base rates are usually one-eighth of one per cent above this.

Where the interest rate is an important element in a major change in economic policy, the Bank will announce it is *not* dealing in the market but that it will lend at 2.30 p.m. at a *stated* rate. However, this procedure, like the retained 'special deposits' weapon, is rarely used today.
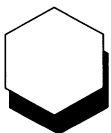
Part VII
Managing the economy

⬡ 15 Making the most of limited resources

15.1 The nature of the problem

(a) The limit to production

Although at any one time resources are limited, they have two important characteristics: (i) they are capable of alternative uses, and (ii) they can be increased over time. The first enables us to make the most of what we have; the second allows us to improve our lot. We can illustrate the situation as follows.

Suppose that, during a year, a country can, with all its resources fully employed, produce the following alternative combinations of agricultural produce and manufactured goods (in unspecified units):

Agricultural produce	+	Manufactured goods
100		0
80		25
60		40
40		45
20		48
0		50

By plotting these alternative combinations we obtain a 'production possibility curve' (Figure 15.1). This shows the various combinations of agricultural
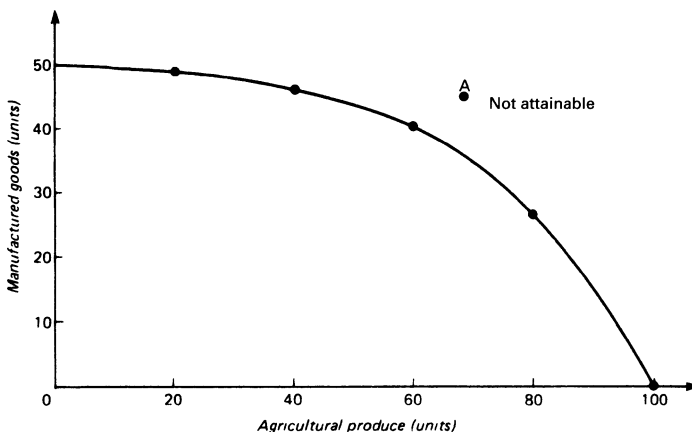
Figure 15.1 A production possibility curve

produce and manufactured goods attainable by country X with its limited resources and given technology.

Nevertheless, as resources are transferred to manufacturing from agriculture, an ever increasing quantity of manufactured goods has to be given up to obtain an extra 20 units of agricultural produce. For instance, when 20 agricultural units are produced, the opportunity cost of an extra 20 is only 5 manufactured goods, whereas when production is 60 agricultural units the opportunity cost of a further 20 is 15 manufactured goods.

The reason is that resources are not equally suited to producing agricultural produce and manufactured goods. For instance, factory workers would need training in farm work, while land, tractors, etc. would have to be worked more intensively. The result is that the production possibility curve is concave to the origin.

The production possibility curve shows that, in order to obtain as high a standard of living as possible, country X must be concerned with:

(i) *The allocation of resources*
Because resources can be used in alternative ways, it is necessary to choose between manufactured goods and agricultural produce. The particular combination chosen should be the one which yields maximum satisfaction.

(ii) *The full employment of resources*
While it is impossible to produce a combination of agricultural produce and manufactured goods which lies to the right of the curve, any assortment on the origin side means that the country is not maximising its output because some resources are idle. A combination of goods on the production possibility curve can be obtained only if there is full employment.

(iii) *The growth of resources*
Over time, the production possibility curve must be pushed further from the origin so that a larger assortment of both agricultural produce and manufactured goods can be obtained.

(b) Micro- and macro-economics
So far we have concentrated mainly on the first problem – the allocation of a given quantity of resources between different uses. By dividing the economy into a number of comparatively small parts, we have been able to study how each part functions – the demand of consumers in a particular market, the behaviour of the firm, the price of the commodity, and so on. Such subjects come within the realm of 'micro-economics'. Thus, if we ask ourselves what forces determine the price of potatoes, the rent of an acre of land in London, the dividend on a particular equity or the wage of a Nottingham bus-driver, we are dealing with micro-economic questions.

We now turn to the second and third problems, full employment and growth. These are concerned with the economy as a whole. They give rise to a series of general questions. How do fluctuations in the overall level of employment occur?

How can overall demand in the economy change? How do firms in general respond to such a change in demand? What brings about changes in the general level of prices? Such questions are the concern of macro-economics.

Referring back to Figure 1.3, micro-economics is concerned with the centre of the diagram – an examination of a particular market and of the connection between one market and another, e.g. those for pencils and pens. But it does not analyse the relationship of any one market to the system as a whole. This is covered by macro-economics, which examines in aggregate the two outer flows, of goods and services and of factors of production.

(c) The method of macro-economics
Because micro- and macro-economics differ in the types of subject they study, their methods of analysis also differ. A simple analogy will explain why.

We can investigate the working of a motor-car by examining its different parts in isolation from one another. Thus we look at the wheels, then the gearbox, the engine, the carburettor, the electric fuel-pump, and so on. In this way we can find out how each part of the car works in detail.

Now, while such an examination is very important and useful, it has its limitations. This is because we just spotlight one component and see how it operates, *ignoring the rest of the car.* It will not enable us to predict what will happen if, for instance, we replace a one-litre engine by a two-litre engine. We cannot assume that, 'other things being equal', the larger engine will make the car run faster. There will be certain 'feedbacks' on the other parts of the car which will affect its running efficiency as a whole. Thus the larger engine may be too powerful for the gearbox; the carburettor may be unable to supply sufficient fuel; the suspension may not be capable of withstanding higher speeds; and so on. It is not enough to examine how one part of the car works in isolation; we have also to consider how the various components are interrelated, and the relative importance of each.

The same applies when we study how an economic system works. The micro-approach will only take us so far, for it merely examines how small parts of the economy operate in isolation. Changes which are simultaneously taking place in other parts of the economy and in the level of activity in general are ignored by inserting the phrase 'other things being equal'. So are the repercussions – the feedbacks – which may result from the single change being analysed.

Now this is legitimate enough if we are analysing a comparatively insignificant part of the economy, for example a small industry. Thus if we wish to discover the effects of an increase in the demand for pencils, we are unlikely to make serious errors by assuming 'other things equal', for such a change is unlikely to cause repercussions on the economy as a whole.

But what if we are examining the effects of a considerable increase in the demand for cars? The car industry is a significant part of the whole economy, and so we cannot merely analyse the effects in isolation, stopping at the point where the price of cars rises or the wages of car workers increase. To indicate the full economic results, we shall have to consider (i) possible 'feedbacks' on the demand for cars, and (ii) repercussions on the economy as a whole – which in turn can produce further feedbacks. For one thing, we shall have to know the

level of employment in the economy. If this is low, the increased demand for cars may lead to a considerable expansion of activity throughout the economy, leading to a further increase in the demand for cars. On the other hand, if the economy is already running at full employment, the increased demand may merely cause higher prices.

(d) Simplifying by aggregating

Although, when dealing with changes in the economy as a whole, we cannot assume 'other things being equal', we still need to simplify if we are to build up a satisfactory model. This we do by 'aggregating' variables into a few broad groups.

   The main aggregates we examine in macro-economics are national income, national output and national expenditure. But we can also deal with sub-aggregates and analyse the factors that determine these. Thus, in analysing national expenditure, we examine consumption spending, investment spending,` government spending, export receipts, spending on imports, etc. Similarly, when looking at national income, we consider wages, rent and profits (see Chapter 16). Aggregating in this way enables us to handle all the different variables so that we can bear in mind the effects which a change in any one of them will have on the other groups and upon the level of activity as a whole (see Chapter 18).

(e) The economic system

What we have said so far must not be taken to imply that particular markets and the economy as a whole are mutually exclusive. To return to our analogy of the car, all the parts of a car are 'ticking over' when the car is running. Each bit of the carburettor and gearbox, and the way each is functioning, affect the overall running of the car. And how the car is driven will influence the performance of the individual parts – the engine, the suspension, the wheels, and so on. So it is with the economy. The millions of independent decisions made by individual firms and consumers affect the overall functioning of the economy. For example, each decision of an individual firm regarding alterations to its factory, plant or stocks affects the amount of investment spending which the economy as a whole is undertaking. And the firm's decisions will also be influenced by the price of its product – which in its turn depends upon the demands of individual purchasers.

   Both micro- and macro-economics are necessary, therefore, to an understanding of how the economy functions. However, before considering the influences on the level of activity, we look at the determinants of the material standard of living of a country (that is, the current position of the production-possibility curve) and the nature of government economic policy.

15.2   Factors determining a country's material standard of living

Since people can enjoy only what they produce with their limited resources, the production-possibility curve shows the limit to their material standard of living. Because income is not evenly distributed, however, what we are talking about is

an average standard of living, usually measured by the national income per head of the population (see Chapter 16).

The factors which limit the standard of living can be classified as internal and external, the latter resulting from economic relationships with the rest of the world.

The most important *internal* factors are:

(a) Original natural resources

Obviously, 'natural resources' cover such things as mineral deposits, sources of fuel and power, climate and the fertility of the soil and fisheries around the coast, but also included are geographical advantages, such as navigable rivers or lakes, which help communications.

While national output increases as new techniques or transport developments allow natural resources to be exploited, the exhaustion of mineral resources works in the opposite direction. Moreover, where a country's economy is pre-dominantly agricultural, variations in weather may cause its output to fluctuate from year to year.

(b) The nature of the people, particularly of the labour force

Other things being equal, the standard of living will be higher the greater the proportion of workers to the total population and the longer their working hours.

But the quality of the labour force is also important. This will depend upon the basic characteristics of the people – their health, energy, adaptability, inventive-ness, judgment and ability to organise themselves and to co-operate in production – together with the skills they have acquired through education and training.

(c) Capital equipment

The effectiveness of natural resources and of labour depends almost entirely upon capital equipment. Thus machinery is necessary to extract oil and minerals, a turbine generator to harness a waterfall, and hotels to exploit Spanish sun and beaches. Similarly, the output of workers varies almost in direct proportion to the capital equipment and power at their disposal. Indeed, the most important single cause of material progress is investment, the addition to capital.

(d) The organisation of resources

To achieve the maximum output from scarce factors of production, they must be organised efficiently. Have we the correct proportion of machinery to each worker? Is the production of the particular good being carried on in the best pos-sible locality? Could the factors be better deployed within the factory? Such questions have to be answered by those organising production.

(e) Knowledge of techniques

Technical knowledge is acquired through capital expenditure on research and invention. Further capital expenditure is necessary to develop discoveries, e.g. to utilise our present knowledge of nuclear energy. Nevertheless the rapid increase of the standard of living of the UK over the last hundred years has largely been

due to the development and application of new inventions such as the steam engine, the internal-combustion engine, electrical power and electronics.

(f) Political organisation

A stable government promotes confidence and thereby encourages saving and investment in long-term capital projects.

To the above we have to add what can be termed *external* factors:

(g) Foreign loans and investments

A net income from foreign investments means that a country obtains goods or services from other countries without having to give goods and services in return, and vice versa. Generally speaking, welfare from this source is only likely to fluctuate over a long period.

(h) The terms of trade

In the short run fluctuations in the terms of trade are likely to be far more important in changing material welfare, especially if the country has, like the UK, a high level of imports and exports.

By the *terms of trade* we mean the quantity of another country's products which a nation gets in exchange for a given quantity of its own products. Thus, if the terms of trade move in a nation's favour, it means that it gets a larger quantity of imports for a given quantity of its own exports. This happens because the prices of the goods that are imported have fallen relative to the prices of those exported. Thus the 1979 increase in the price of oil reduced the standard of living of the importing countries and raised that of the oil producers.

(i) *Gifts from abroad*

Aid to countries for purposes of economic development and defence improve the standard of living of the receiving countries.

15.3   The nature of government economic policy

(a) Objectives of government management of the economy

In chapter 1 we drew together the various reasons for government interference in the market economy, classifying them under three broad objectives: (i) allocation of resources, (ii) stabilisation and (iii) re-distribution of wealth and income.

The first objective – improving the allocation of resources – has already been discussed with reference to market failure. Thus the government, may interfere in the market e.g. by carrying stockpiles to be used to stabilise the prices of certain agricultural products, by controlling monopolies, by allowing for externalities, and by itself providing community goods and services.

In the remainder of this book we examine stabilisation policy and, within the broader context of public finance, the re-distribution of wealth and income.

(b) Government stabilisation policy

Government stabilisation policy is concerned with:

  (i) Full employment.
 (ii) A stable price level.
(iii) Balanced regional development.
(iv) Growth of national income.
 (v) A healthy overseas trading position.

In considering these objectives we have to use macro-economic analysis, looking at groups as a whole – consumers, firms, trade unions, etc. – to see how they are connected in the whole economic system.

Once again it must be pointed out that these policies may not be fully compatible, so that there has to be some 'trade-off' between them. The late Harold Macmillan, a former prime minister, likened the role of the government to that of a juggler whose task is to keep many balls in the air at the same time. Some balls are going up because they have been given special attention, while others are on their way down.

In the last resort the emphasis put on different policies is a matter of judgment and therefore rests on political views. Our task is to analyse possible measures and suggest their likely consequences. The politician then has to decide where the balance of advantage lies.

16 Measuring the level of activity: national-income calculations

16.1 The principle of national-income calculations

Fluctuations in the level of activity are monitored by quantitative information on the national income. Although the collection of statistics proceeds continuously, the principal figures are published annually in *The United Kingdom National Accounts* (*The CSO Blue Book*).

The principle of calculating national income is as follows. Income is a flow of goods and services over time: if our income rises, we can enjoy more goods and services. But for goods to be enjoyed they must first be produced. A nation's income over a period, then, is basically the same as its output over a period. Thus, as a first approach, we can say that national income is the total money value of all goods and services produced by a country during the year. The question is how we can measure this money value.

We can tackle the problem by studying the different ways in which we can arrive at the value of a table.

Figure 16.1 shows that the value of the table can be obtained by taking the value of the final product (£100) or by totalling the value added by each firm in the different stages of production. The output of the tree-grower is what he receives for the tree (£30) which, we will assume, cost £20 in wages to produce, leaving £10 profit. The output of the sawmiller is what he receives for the timber (£50) less what he paid for the tree. Again, this output (£20) is made up of wages and profit. And so on. The total of these added values equals the value of the final table. Thus we could obtain the value of the table by adding the *net outputs* of the
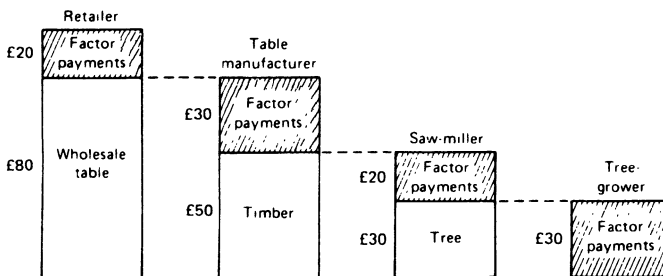
Figure 16.1 The value of the total product equals the sum of values added by each firm

tree-grower (forestry), the sawmiller and the table-maker (manufacturing) and the retailer (distribution).

Alternatively, instead of putting these individual outputs in industry categories, we could have added them according to the type of factor payment – wages, salaries, rent or profit. This gives us the *income* method of measuring output.

Thus, if we assume (i) no government taxation or spending and (ii) no economic connections with the outside world, we can obtain the national income either by totalling the value of final output during the year (i.e. the total of the value added to the goods and services by each firm) or by totalling the various factor payments during the year – wages, rent and profit.

There is, however, a third method of calculating the national income. The value of the table in Figure 16.1 is what was spent on it. If the table had sold for only £90, that would have been the value of the final output, with the final factor payment – profit to the retailer – reduced to £10. Thus we can obtain the national income by totalling *expenditure* on final products over the year.

It must be emphasised that the money values of output, income and expenditure are *identical by definition*. They simply *measure* the national income in different ways. This was shown by the fact that factor payments were automatically reduced by £10 when the table sold for £90 instead of £100.

Before we proceed to examine in more detail the actual process of measuring these three identities, it is convenient if we first consider some of the inherent difficulties.

16.2   National-income calculations in practice

(a) General difficulties
Complications arise through:

(i) *Arbitrary definitions*

(1) PRODUCTION
In calculating the national income, only those goods and services which are paid for are normally included. Because calculations have to be made in money terms, the inclusion of other goods and services would involve imputing a value to them. But where would you draw the line? If you give a value to jobs which a person does for himself – growing vegetables in the garden or cleaning the car – then why not include shaving himself, driving to work, and so on? On the other hand, excluding such jobs distorts national-income figures, for, as an economy becomes more dependent on exchanges, the income figure increases although there has been no addition to real output? (see p. 219).

An *imputed* money value is included for certain payments in kind which are recognised as a regular part of a person's income earnings, e.g. food, etc. provided for the armed forces.

(2) THE VALUE OF THE SERVICES RENDERED BY CONSUMER DURABLE GOODS

A TV set, dishwater, car, etc., render services for many years. But where would we stop if we imputed a value to such services? A toothbrush, pots and pans, for example, all render services over their lives. All consumer durable goods are therefore included at their price when bought, subsequent services being ignored.

The one exception is *owner-occupied houses*. These are given a notional rent to keep them in line with property owned for letting, whose rents are included, either directly or as profits of property companies. This also prevents national income falling as more people become owner-occupiers!

(3) GOVERNMENT SERVICES

Education and health services, although provided by the state, are no different from similar services for which some persons pay. Consequently they are included in national income at cost. But what of certain other government services? A policeman, for instance, when helping children to cross the road is providing a consumer service. But at night his chief task may be guarding banks and factories, and in doing so he is really furthering the productive process. To avoid double-counting, this part ought to be excluded from output calculations. In practice, however, it would be impossible to differentiate between the two activities, and so all the policeman's services – indeed all government services (including defence) – are included at cost in calculating national output (see p. 219).

(ii) *Inadequate information*

The sources from which data are obtained were not specifically designed for national-income calculations. For instance, the Census of Production and the Census of Distribution are only taken at approximately five-year intervals. As a result many figures are estimates based on samples.

Information, too, may be incomplete. Thus not only do income-tax returns fail to cover the small-income groups, but they err on the side of understatement.

But it is 'depreciation' which presents the major problem, for what firms show in their profit-and-loss accounts is affected by tax regulations. Since there is no accurate assessment of real depreciation, it is now usual to refer to gross national product (GNP) rather than to national income (see Figure 16.2).

(iii) *The danger of double-counting*

Care must be taken to exclude transfer incomes when adding up national income (see p. 216), the contribution to production of intermediary firms when calculating national output (see Table 16.1) and indirect taxes when measuring national expenditure (see p. 217).

A fourth way in which a form of double-counting can occur is through 'stock appreciation'. Inflation increases the value of stocks, but although this adds to firms' profits it represents no increase in real income. Such gains must therefore be deducted from the income figure.
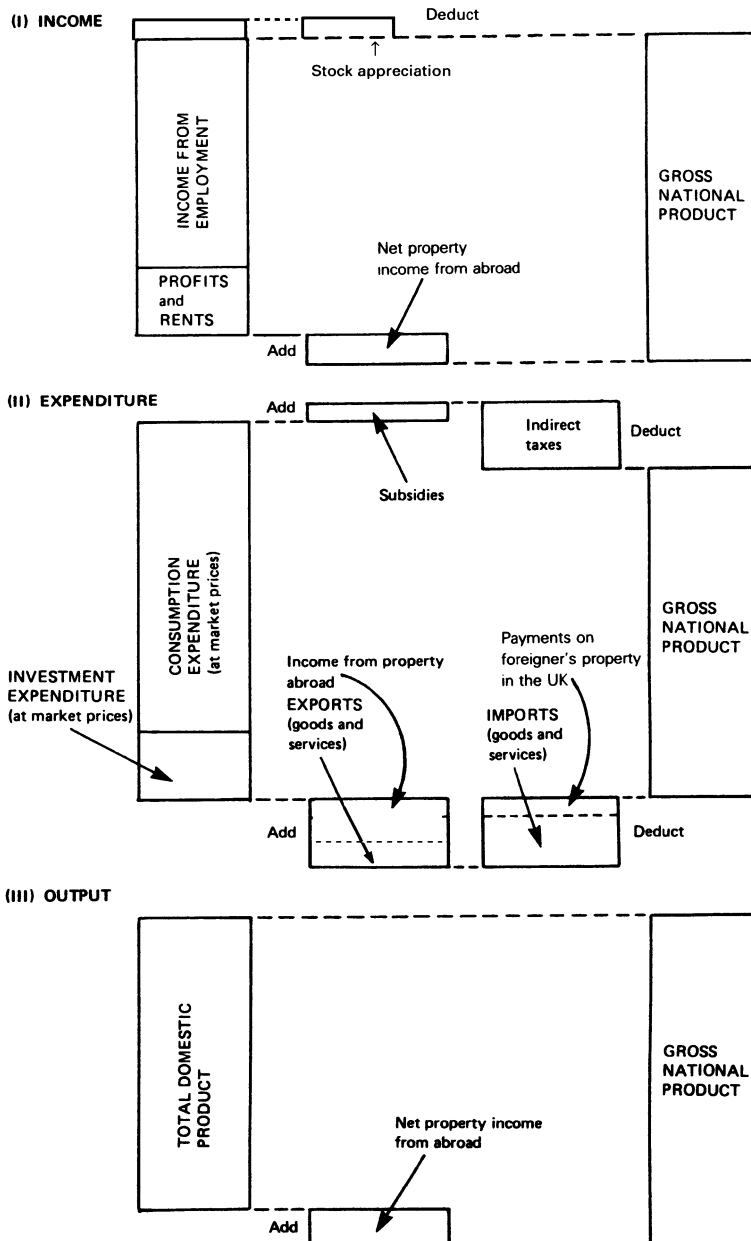
Figure 16.2 Summary of gross-national-product calculations

Table 16.1   Calculations of the national income of the UK, 1991

INCOME	£m
Income from employment	329 808
Income from self-employment	57 505
Gross trading profits of companies	60 674
Gross trading surplus, public corporations	3 119
Gross trading surplus, general government enterprises	119
Rent	44 092
Imputed charge for consumption of non-trading capital	4 490
Total domestic income	499 809
less stock appreciation	–2 825
Statistical discrepancy	17
GROSS DOMESTIC PRODUCT (at *factor cost*)	497 001
Net property income from abroad	328
GROSS NATIONAL PRODUCT	497 329
less capital consumption	– 63 968
NATIONAL INCOME	433 361

EXPENDITURE	£m
Consumers' expenditure	367 853
General government final consumption	121 899
Gross domestic fixed capital formation	95 442
Value of physical increase in stocks and works in progress	–5 303
Total domestic expenditure	579 891
Export of goods and services	135 115
Total final expenditure	715 006
less imports of goods and services	–140 415
Statistical discrepancy	– 445
GROSS DOMESTIC PRODUCT (at *market prices*)	574 146
less taxes on expenditure	– 83 023
plus subsidies	5 878
GROSS DOMESTIC PRODUCT (at *factor cost*)	497 001
Net property income from abroad	328
GROSS NATIONAL PRODUCT	497 329
less capital consumption	– 63 968
NATIONAL INCOME	433 361

Table 16.1   cont'd.

OUTPUT	£m
Agriculture, forestry, and fishing	8 772
Energy and water supply	28 273
Manufacturing	104 283
Construction	33 686
Distribution, hotels and catering: repairs	73 024
Transport and communication	34 755
Banking, finance, insurance, business services and leasing	88 179
Ownership of dwellings	34 839
Public administration, national defence and compulsory social security	34 786
Education and health services	49 643
Other services	33 915
Total	524 155
Adjustment for financial services	–27 171
Statistical discrepancy (income adjustment)	17
GROSS DOMESTIC PRODUCT (income based)	497 001
Net property income from abroad	328
GROSS NATIONAL PRODUCT	497 329
less capital consumption	– 63 968
NATIONAL INCOME	433 361

(iv) *Relationship with other countries*

(1) TRADE
British people spend on foreign goods, while foreigners buy British goods. In calculating national *expenditure,* therefore, we have to deduct the value of goods and services imported (since they have not been produced by Britain) and add the value of goods and services exported (where income has been earned in Britain).

(2) INTERNATIONAL INDEBTEDNESS
If a father increases his son's pocket-money, it does not increase the family income. Instead, it merely achieves a re-distribution, the father having less and the son more. But if the boy's aunt makes him a regular allowance, the family income is increased. Similarly, with the nation: while transfer incomes, e.g. retirement pensions and student grants, do not increase national income, pay-ments by foreigners do. These payments arise chiefly as interest and dividends from loans and investments made abroad. In the same way, foreigners receive payments for investments in Britain. Net income from abroad (receipts less payments) must therefore be added to both domestic expenditure and output.

(b) Government calculations of the national income

We start off by measuring Gross *Domestic* Product (GDP). The GDP is simply the money value of the final output of all resources located within a country irrespective of whether their owners live there or abroad. Hence in order to obtain Gross *National* Product (GNP) we have to add the balance of *net* property income from abroad (Figure 16.2).

Figures for GNP are calculated for income, expenditure and output. Because information is incomplete, results from the three methods vary. Thus an 'average figure' is built up and the difference from each calculation is shown as a 'statistical discrepancy'.

(i) *National income*

National income is the total money value of all incomes received by persons and enterprises in the country during the year. Such incomes may be in the form of wages, salaries, rent, or profit.

In practice income figures are obtained mostly from income-tax returns, but estimates are necessary for small incomes. Two major adjustments have to be made:

(1) TRANSFER INCOMES

Sometimes an income is received although there has been no corresponding contribution to the output of goods and services, e.g. unemployment-insurance benefit and interest on the National Debt. Such incomes are really only a re-distribution of income within the nation – chiefly from taxpayers to the recipients. Transfer incomes must therefore be deducted from the total of all incomes.

(2) INCOME FROM GOVERNMENT ACTIVITIES

Personal incomes and the profits of companies are obtained from tax returns.

Trading activities of public corporations, e.g. the Post Office and the BBC, and of local authorities, e.g. housing, transport, may also show surpluses which have to be added in, while an imputed rental value in given to the property owned and occupied by the government and local authorities (non-trading income).

(ii) *National expenditure*

National expenditure is the total amount spent on consumer goods and services and on net additions to capital goods and stocks in the course of the year.

Figures for calculating national expenditure are obtained from a variety of sources. The Census of Distribution records the value of shop sales, while the Census of Production gives the value of capital goods produced and additions to stocks. But these censuses are not taken every year, and gaps are filled by estimates from data provided by the National Food Survey and the Family Expenditure Survey.

Market prices are swollen by indirect taxes, e.g. VAT, and reduced by subsidies, e.g. on council housing. What we are trying to measure is the value of the

national expenditure which corresponds to the cost of the factors of production (including profits) used in producing the national product. This is known as 'national expenditure at factor cost' and is obtained by deducting indirect taxes from and adding subsidies to national expenditure at market prices.

Adjustments necessary for exports and imports have already been referred to (see p. 215).

(iii) *National output*

National output is the total of consumer goods and services and investment goods (including additions to stocks) produced by the country during the year. It can be measured by totalling either the value of the *final* goods and services produced or the *value added* to the goods and services by each firm, including the government.

(c) Gross national product and national income

In the course of production, machinery wears out and stocks are used up. This represents depreciation of capital. If we make no allowance for this but simply add in the value of new investment goods produced, we have *gross national product*. But, to be accurate, the calculation of total output should include only net investment – that is, the value of new investment goods and stocks less depreciation on existing capital and stocks used up. This gives the net national product, which is the true national income for the year (Figure 16.3).
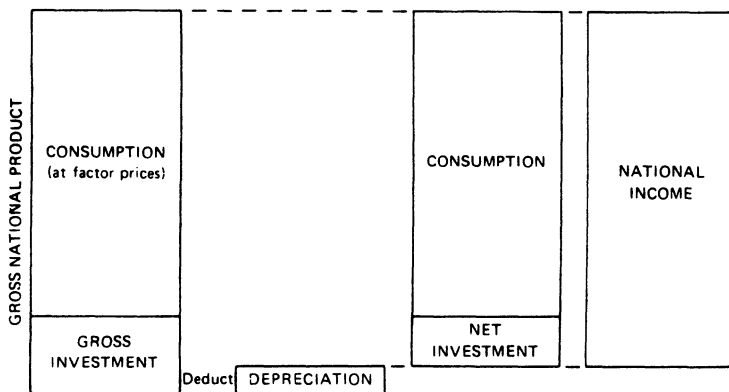
Figure 16.3 Gross national product and national income

(d) Personal disposable income

For some purposes, e.g. as an indication of people's current living standards, a measurement of personal disposable income, that is, what people actually have to spend, is more significant. The necessary adjustments to gross national product to obtain personal disposable income are shown in Figure 16.4.

Net income
from foreign assets

Income tax
Company taxes
National Insurance Contributions

Deduct

Profit and rents
of public
undertakings

GROSS NATIONAL PRODUCT

WAGES AND
SALARIES

PROFIT AND
RENT

Deduct

Deduct

Undistributed
profit and depreciation

Deduct

PERSONAL
DISPOSABLE
INCOME

Add | GOVERNMENT
TRANSFER
PAYMENTS
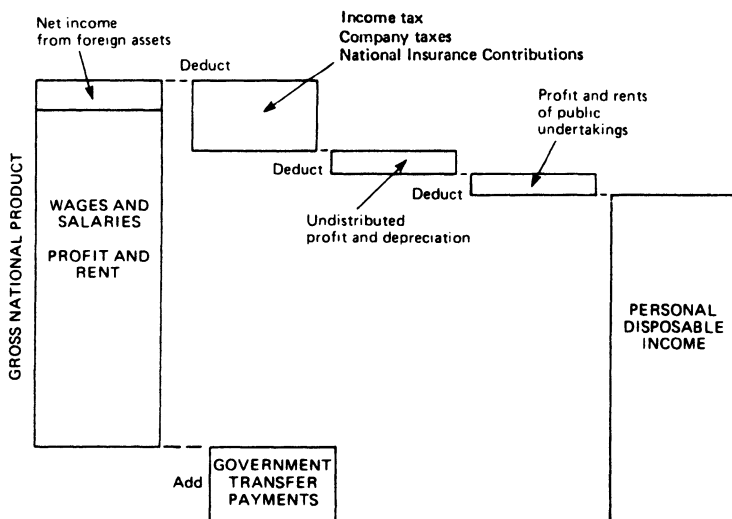
**Figure 16.4   The relationship between gross national product and personal
disposable income**

16.3   Uses of national-income calculations

(a) To indicate the overall standard of living

Welfare is not identical with wealth (see p. 10), but wealth bears the closest sin-
gle relationship to it. Income, the flow of wealth, is therefore the nearest indica-
tion of welfare.

Nevertheless, the national-income figure cannot be accepted solely at its face
value. Thus, although the national income of the UK was £227 764 million in
1983 and £433 361 million in 1991, it does not automatically follow that
everybody had doubled their standard of living over that period. The following
qualifications have to be made:

  (i) Some of the increase may be due to inflation (see Chapter 19).
 (ii) The national-income figure must be related to the size of population;
     thus average income per head is a better indication of well-being.
(iii) A person's standard of living depends upon the quantity of consumer
     goods and services he enjoys. But the increase in national income may
     have come about mainly through an increase in the production of pro-
     ducer goods. While these goods enable a higher standard of living to be
     enjoyed in the future, they do not increase *present* welfare. Thus average
     personal disposable income might provide a better indication of cur-
     rent living-standards, though national income per head is the more
     satisfactory in the long run.
 (iv) The increase in national income may have come about through a surplus
     of exports over imports. This represents investment overseas, and thus
     (iii) above applies.

(v) The average-income-per-head figure is merely a statistical average. It does not indicate how any increase in national income is distributed; it may go mostly to a few rich people (as in the oil sheikhdoms), perhaps leaving the others little better off.

(vi) National-income figures do not reflect the 'quality' of life. An increase in national income may be the result of longer working hours, inferior working conditions, longer journeys to work, or the presence of more housewives at work (with less comfort in the home). The national income does not value leisure time.

(vii) The quality of goods may have improved without any increase in real price, e.g. razors, biros.

(viii) All government spending is included at cost in national-income calculations, no distinction being made between expenditure on consumer services and expenditure on defence. As a result, if spending on the social services were cut to pay for re-armament, national income would be unchanged!

(ix) The national-income figure is swollen when people pay for services which they previously performed themselves. Thus a married woman who returns to teaching but pays a woman to do her housework adds to the national income twice – although only the net addition is her teaching services.

(x) Because national-income figures are based on private costs and benefits, external costs or benefits do not enter into the calculations. Thus the erection of electricity pylons would be included at cost, no allowance being made for the social cost of spoiling the landscape.

(xi) The increase in national income may have incurred an excessive consumption of irreplaceable resources, e.g. fossil fuels.

(xii) National income figures do not include the black economy (some estimates would add 10 per cent) where services are exchanged for cash in order to evade taxes.

(xiii) National income includes payments necessitated by the stress of modern living, e.g. for anti-depressant drugs, clinics for alcoholics.

(b) To compare the standards of living of different countries

Comparisons of the national incomes of different countries are often necessary for practical purposes. How much help should be given by the rich countries to the very poor? Which are the very poor countries? What contribution should be made by a country to an international body, such as the United Nations, the EC and NATO? What is the war potential of a country as indicated by its GNP per head?

But, when used to compare countries' standards of living, national income figures must be subjected to qualifications additional to those mentioned in (a) above.

(i) Because figures are expressed in different currencies they have to be converted into a common denominator. Using the exchange rate for this purpose is not entirely satisfactory, for the rate is determined by factors other than the internal purchasing powers of currencies, e.g. capital movements. More satisfactory is the use of a conversion rate based on the purchasing power parity for a common basket of goods.

(ii) Different people have different needs. The Englishman has to spend more on heating than the Indian. Obviously, the Englishman is no better off in this respect – though the national-income figures, by valuing goods at cost, would indicate that he is.

(iii) The proportionate of national income spent by different countries on defence varies. Countries which spend less can enjoy consumer goods instead, but average national income does not indicate the difference.

(iv) Countries vary as regards the length of the average working week, the proportion of women who work, the number of jobs which people do for themselves, the degree to which goods are exchanged against money, the size of the 'black economy' and the accuracy of tax returns. Some allowance must be made for each of these factors.

(c) To calculate the rate at which a nation's income is growing

Is the national income growing? Is it growing as fast as it should? Are the incomes of other countries growing faster? Is there sufficient investment to maintain future living-standards? The answers to these and similar questions can be found by comparing national-income figures, though for the reasons given above some caution must be observed.

(d) To establish relationships which arise between various parts of the economy

If, for example, national-income figures revealed a relationship between the level of investment and growth, or between educational expenditure and growth, or between profits and the level of investment, such information would be useful in planning the economy.

The figures might also indicate trends, e.g. the proportion of national income which is taken by the government.

(e) To assist the government in managing the economy

Some central-government planning is now regarded as essential for achieving full employment, a stable currency and a satisfactory rate of growth. But this requires having figures for the various components of the national income, such as consumption spending, investment, exports and imports. How they can be used will be explained later.

(f) To assist businesses, trade unions, financial journalists, etc. to ascertain economic trends and forecast future movements

(g) To indicate changes in the distribution of income

While, as a scientist, the economist is not concerned with the 'fairness' of the distribution of income, the government is, for taxation and political considerations. National income figures provide the statistical basis when deciding on such matters.

⬡17 Unemployment and its causes

17.1 What do we mean by 'full employment'?

Although today 10% of the working population are unemployed, this compares favourably with the situation in pre-war Britain, where, in the worst year – 1932 – the national unemployment rate was 22.1%. Unemployment means that labour, machines, land and buildings stand idle; as a result, the standard of living is lower than it need be. But the real curse is the human misery that results. Many people, without work for years, lose hope of ever finding a job; in any case skills deteriorate as the period of unemployment lengthens. Thus unemployment is usually discussed in terms of labour.

    We say that unemployment exists where people capable of and willing to work are unable to find suitable paid employment, but official figures include only those actually claiming benefit. Where an economy is constantly adapting to changing conditions, there will always be some persons unemployed as they switch jobs or as seasonal or casual works comes to an end. Some 3% unemployment must be allowed for, to cover this. More important, trying to run the economy at a very low level of unemployment gives rise to problems – inflation, pressure on the balance of payments, a lack of competitiveness and a slow rate of growth – which eventually thwart the objective. Thus the government has to weigh the pros and cons of such a policy. The decision is eventually a political one, but experience since 1970 suggests that any government, whatever its political colour, will have difficulty in achieving less than 5 per cent unemployment.

17.2 Causes of unemployment

Unemployment may occur for many different reasons, and these must be distinguished if the appropriate remedies are to be applied.

(a) *Frictional unemployment*
Unless the economy is completely static, there will always be people changing their jobs. Some merely desire a change of employment or a move to a different part of the country. In certain occupations, e.g. unskilled labour in the construction industry, workers are not employed regularly by any one employer: when a particular contract is completed, labour is made redundant. Occasionally, too, workers are discharged when a factory is being reorganised.

    Unemployed workers usually register at the local Jobcentre, forming a pool of labour from which employers can fill vacancies. But how large should this pool be? If it is too large, workers remain unemployed for long periods. If it

is too small, production is dislocated by bottlenecks in filling vacancies (with employers holding on to labour not currently needed), by job-switching just for the sake of change and, above all, by strikes in support of claims for higher wages.

(b) *Seasonal unemployment*
Employment in some industries, e.g. building, fruit-picking and holiday catering, is seasonal in character. The difficulty is that the skills required by different seasonal jobs are not 'substitutable'. To what extent, for example, can hotel workers become shop assistants in the January sales? Seasonal employment is not completely avoidable. But it can be reduced if a small, regular labour force will work overtime during the 'season' and admit such people as students and housewives during the busy periods. Moreover, the price system may help. By offering off-season rates, hotels at holiday resorts can attract autumn conferences.

(c) *International causes*
Because the UK is so dependent on international trade, she is particularly vulnerable to unemployment brought about by a fall in the demand for her exports. Such a fall may occur because:

(1)   THE PRICES OF UK GOODS ARE TOO HIGH TO BE COMPETITIVE IN WORLD MARKETS
If home prices rise, for example because of wage increases, the export market is likely to be hit severely. The demand for exports is usually highly elastic, since substitutes are often available from competing countries.

Quantity of export good demanded and supplied
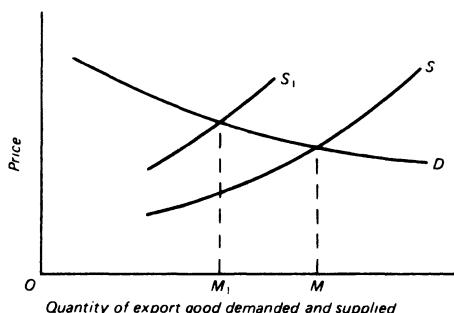
Figure 17.1   The effect on employment of a wage increase in an export industry

The effect on employment is shown in Figure 17.1. The wage increase moves the supply curve from S to S_1. Because demand is elastic there is a considerable fall in the demand for the good, from OM to OM_1. The industry, and therefore employment, contract.

(2) INCOMES OF MAJOR IMPORTING COUNTRIES MAY BE REDUCED BY A RECESSION OR A
    DETERIORATION IN THE TERMS OF TRADE

If incomes of importing countries fall, their demand for UK goods, especially those having a high income-elasticity of demand, will be likely to decrease. This happened during the world recession 1993.

(d) *Structural unemployment*

Structural unemployment, like the frictional type, results largely from the immobility of labour; but in this case it is brought about by long-term changes in the conditions of demand and supply. It is associated, therefore, with major changes in the economy, particularly in the export industries.

On the demand side, there may be a change in any of the factors influencing the conditions of demand. The price of substitutes may fall (coal has largely been replaced by oil and gas), or foreign buyers may switch to competitors' goods (British shipyards have been hit by Japanese competition). On the supply side, new techniques or the exhaustion of mineral deposits may make labour redundant. Thus natural gas has led to the closure of coal mines.

Where an industry is highly localised in a particular area, the resulting unemployment may be particularly severe (see p. 253)

The above types of unemployment are largely the result of frictions – the immobility of labour – which make the price system work imperfectly. Here an approach based on demand and supply in a *particular* labour market can provide an explanation (chapter 10).

(e) *Cyclical unemployment*

The term 'cyclical unemployment' refers to the alternate booms and slumps in the level of industrial activity which have occurred over the last hundred years. It was the major cause of the high unemployment of the 1930s, and while we no longer speak of a 'trade cycle', we do still move between boom and recession.

Policies to deal with cyclical unemployment are more complex. Their nature, and the difficulties involved, are analysed in the chapters which follow.

⬡ 18 The level of output and aggregate demand; defects of demand management as a policy for full employment

18.1   The link between spending and production

(a) The circular flow of income
We will begin by repeating in simplified form the identity which exists between income and expenditure. Take a simple example. A teacher buys a table from a carpenter. With the money he receives, the carpenter pays the timber merchant for the wood, who in turn pays the man who cut the wood. But where did the teacher obtain the original money to buy the table? Simply from the carpenter, the timber merchant and the tree-feller, who each use part of their receipts to pay fees to the teacher for instructing their children. So with the other goods the teacher buys. Thus there is a circular flow of income – one person's spending becomes another person's income. Spending is therefore necessary for earnings.

The same applies to the economy as a whole; at any one time spending equals income. Suppose, for instance, that all production in the economy is in the hands of a giant firm which owns all the land and raw materials and employs all the labour. The firm's income consists of the receipts from the sale of its product. Since it owns all the raw materials and land, these receipts must equal what it pays out in wages and what it has left in profits. This was the principle upon which we measured national income.

Since spending on goods determines the receipts and thus the profits of firms, it is of vital importance in deciding the level of their output and thus of the aggregate level of activity. To explain more fully, we use Figure 18.1, which shows the money flows which correspond to the movement of factors and goods in the outer ring of Figure 1.3 – payments by firms for factors and expenditure of households on goods. The first represents income of households; the second represents receipts of firms.

If spending on goods and services is maintained, factor payments can be maintained; in other words the profitability of production is unchanged and thus firms have no cause to vary output. If, however, for some reason or another, spending should fall, some of the goods produced by firms will not be sold, and stocks will accumulate. On the other hand, if spending on goods and services increases, stocks will be run down. Production has become more profitable and, as a result, output is expanded.

Three important points emerge from our discussion so far:
(i)   There is no impetus towards a contraction or expansion of production if spending on goods and services equals spending (including normal profits) by firms on factors of production.
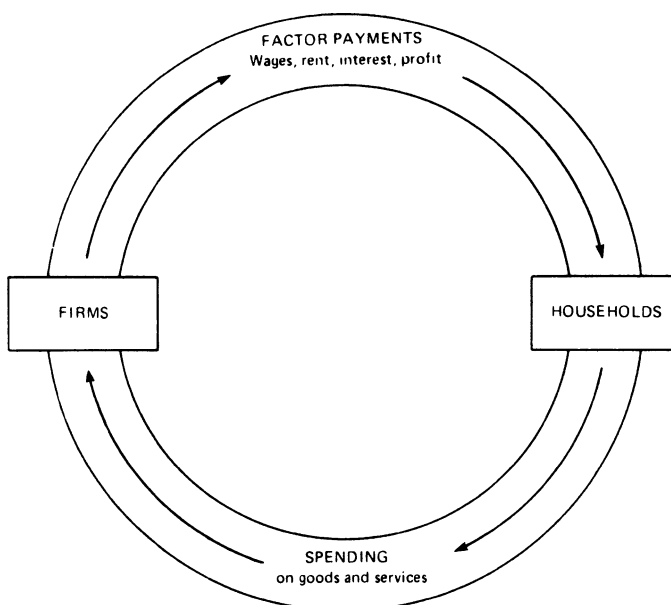
Figure 18.1 The circular flow of income

  (ii) The level of production, and therefore employment, is closely related
       to the level of spending.
 (iii) There is nothing to guarantee that spending will be sufficient to ensure
       a level of production where all factors are fully employed.

Definitions and assumptions
Before we show how changes in spending occur, we must make simplifying
assumptions.
   We define net profit as gross profit less retentions for depreciation. We assume:
(i) *All retentions for depreciation are actually spent on replacement investment*
Thus, when in future we speak of 'investment' it refers solely to net additions to
fixed capital and stocks, i.e. net investment.

(ii) *All net profit is distributed to the owners of the risk capital*
If we define net profit as being gross profit less retentions for actual spending on
depreciation, this means that there is no 'saving' by firms.

(iii) *There is no government taxation or spending*

(iv) *There are no economic connections with the outside world; it is a 'closed'*
     economy
From the above assumptions it follows that: (1) the sum of the factor payments is
equal to national income (equals national output) as defined in chapter 15; and
(2) income equals disposable income.

(v) *There are no changes in the price level*
Thus any changes in the money value of national income reflect changes in real output.

(vi) *The level of employment is directly proportionate to the level of output*
In practice this may not be strictly true: existing machinery, for example, may be able to produce extra output without additional labour. But until there is full employment, the simplification does allow employment to vary directly with the level of national income.

18.2 Reasons for changes in aggregate demand

(a) Aggregate demand
Our task, therefore, is to discover why changes occur in the national income (hereafter symbolised by Y). Now, as we have just shown, Y depends upon the level of spending, which we shall refer to as aggregate demand (abbreviated to AD). Thus we can find out why Y changes by discovering why AD changes.

(b) Changes in AD
Let us return to our example of the teacher. Suppose he earns £16,000 in a given year. Most of it will be spent on consumer goods and services – but not all. Some will probably be put aside for a 'rainy day'. That part of income which is not spent we can say is 'saved'. What happens to it? The money could be hidden under the mattress; in that case it is obviously lost to the circular flow of income. But the teacher is more likely to put it in a bank, where it is safer and earns interest. Nevertheless, at this point it is still lost to the circular flow. Saving represents a 'leak' from the flow of income.

So far, however, we have looked only at spending on consumer goods. But spending may also be on capital goods. Firms borrow money from their banks (and other institutions) for such purchases. Thus the sum deposited by the teacher stands a good chance of being returned to the circular flow of income by being 'invested', i.e. spent on capital goods or additions to stocks. Investment, therefore, can be regarded as an 'injection'. And, if exactly the amount of money saved by households is spent by firms on investment, the level of AD is maintained (Figure 18.2), and Y is unchanged.

But suppose that the amount saved does not coincide with what firms wish to invest. This can come about either by a change in the amount invested or by a change in the amount spent by consumers.

Let us first assume that households' spending on consumer goods remains constant. If now firms reduce the amount they borrow for investment, AD is smaller. On the other hand, if firms increase their investment, AD will be larger.

Alternatively the amount of income spent on consumer goods may alter. Investment, we will now assume, remains unchanged. Here, if more is spent out of a given income, AD will increase; if less, AD decreases.

What is important to recognise is that in an economy where people are free to dispose of their income as they please, and where firms make their own invest-
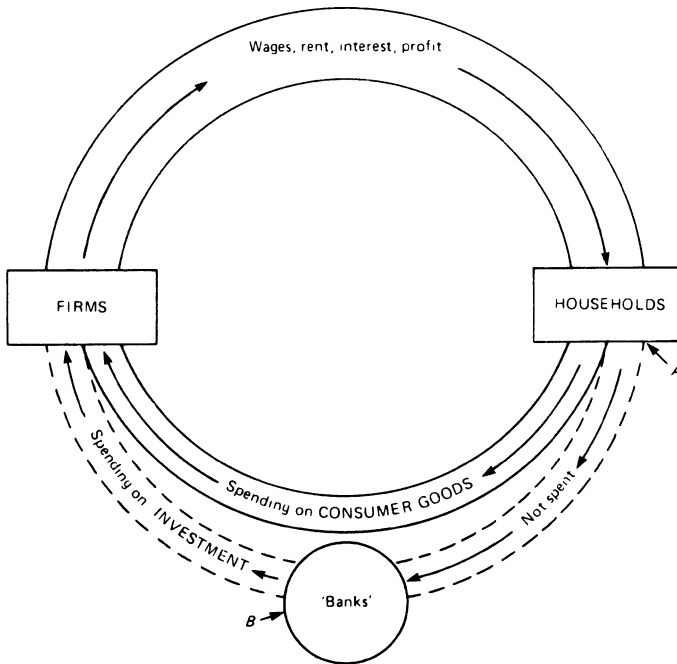
Figure 18.2 The level of income maintained through investment

ment decisions, a difference can easily exist between the amount of income which people plan to 'save' (i.e. which they do not wish to spend) and the amount which firms wish to invest. This is because, in their spending, households and firms act for different reasons. Two questions have therefore to be asked: (i) What determines spending on consumer goods and therefore saving (at position A)? (ii) What determines investment spending (at position B)?

In our analysis, consumption, that is spending on consumer goods and services, will be given the symbol C; saving, i.e. income not spent on consumption, S; investment, i.e. spending on net additions to capital goods and stocks, I.

18.3   Consumption spending

(a) Consumption and saving by households: 'personal saving'
Income is received as wages or salaries, rent, interest and profits. With it households buy the consumer goods they need. That part of income which is not spent has been defined as 'saving'. Hence $Y = C + S$, $C = Y - S$, and $S = Y - C$.

C and S, therefore, are merely two sides of the same coin. Thus, whenever we consider C or S, we must examine the factors which influence both spending and thrift.

Spending decisions are more important in the short run, for a people's first concern is to maintain their standard of living. They are influenced by:

(i) *Size of income*

A small income leaves no margin for saving. Only when basic needs have been satisfied will a part of income be saved. Indeed, if current income falls below this level, past savings or borrowing may be used to maintain the standard of living accustomed to.

But we can go further. As income increases, the proportion spent tends to decrease; or, as it is often put, there is a *diminishing marginal propensity to consume (c)*.

The above conclusions are illustrated diagrammatically in Figure 18.3, where the curve *C* shows how consumption changes with income.

Below an income of *OD* there is 'dis-saving'. At *OD* all income is consumed. At higher incomes the proportion spent falls and saving occurs. This diminishing marginal propensity to consume is shown by the decreasing slope of the consumption curve: for any given increase in income, the extra amount spent grows successively smaller.
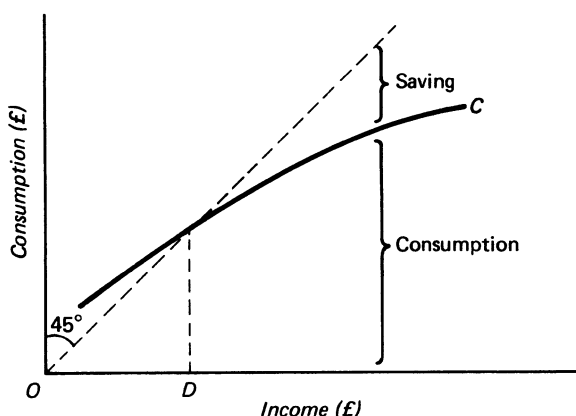
Figure 18.3   The relationship between consumption and income

(ii) *The timelag in adjusting spending habits*

It takes time for a person to adjust his standard of living as his income increases. In the short period, therefore, saving increases.

The above two factors explain the *shape* of the consumption curve – how spending changes as income changes. But we still have to account for the position of the curve – what determines the proportion of any given income which is spent. This amount can vary (that is, the position of the *C* curve may change) as a result of:

(iii) *Changes in disposable income*

We have assumed that firms have distributed all net profits and that there are no government taxation or transfers. In practice both profit distribution and govern-

ment taxation will affect the size of income available for spending. Increasing direct taxation, for instance, would, by reducing disposable income, lower the *C* curve.

(iv) *The size of wealth owned by an individual*

(v) *The invention of new consumer goods*
In recent years family cars, TV sets, hi-fi equipment, camcorders, central heating and dishwaters have all induced spending, especially when backed by intensive advertising.

(vi) *Hire-purchase and other credit facilities*
Easier hire purchase or bank credit terms encourage spending.

(vii) *An anticipated fall in the value of money*
If people considered that the prices of goods were likely to rise, they would bring forward their spending rather than save for the future.

(viii) *Inflation*
Uncertainty regarding possible government measures to reduce inflation may lend to increased saving. Furthermore, inflation reduces the real value of assets fixed in money terms. People may be induced to save more to restore the real value of their assets, e.g. 1990–2.

(ix) *The age-distribution of the population*
Since most saving is done by people over thirty-five years of age, an ageing population will tend to reduce the propensity to consume of the community as a whole.
    In the long period, people have some concern for their future standard of living, and *thrift* exercises a greater influence on the disposal of income.
    The main *factors determining thrift* are:

(i) *Size of income*
As already shown, saving increases as income increases and at an increasing rate.

(ii) *The life cycle*

(iii) *Psychological attitudes*
Some communities are by nature more thrifty than others, providing against sickness, unemployment and old age. On the other hand, ostentation – the desire to 'keep up with the Joneses' – may motivate a high rate of spending.

(iv) *Social environment*
Apart from influencing the general attitude to saving, environment can be a major factor in other ways. Such institutions as savings banks, building societies, insurance companies and unit trusts encourage regular saving.

Political conditions, too, influence saving habits. Countries continually threatened by war or revolution do not provide the stable background necessary to encourage thrift.

(v) *Government policy*

The government can influence people's attitude to saving in a variety of ways. In the UK it tries to stimulate personal saving through the rate of interest offered, income-tax concessions (e.g. on National Savings Bank interest) and special devices (e.g. Personal Equity Plans, Premium Bonds). On the other hand, a comprehensive state social-insurance and pension scheme may reduce personal saving.

At one time it was thought that people could only be induced to postpone consumption, i.e. to save, by offering interest as compensation. This view, however, is now largely rejected, chiefly because much saving is contractual, e.g. pension and mortgage payments. The dominant factor is the *ability* to save, i.e. the level of income.

Under our simplifying assumptions that all net profits are distributed and that there is no government taxation or spending, all saving is done by households. But saving can be achieved through retentions by firms and the government (Figure 18.4). In order to consider these, we will temporarily relax assumptions (2) and (3) above.
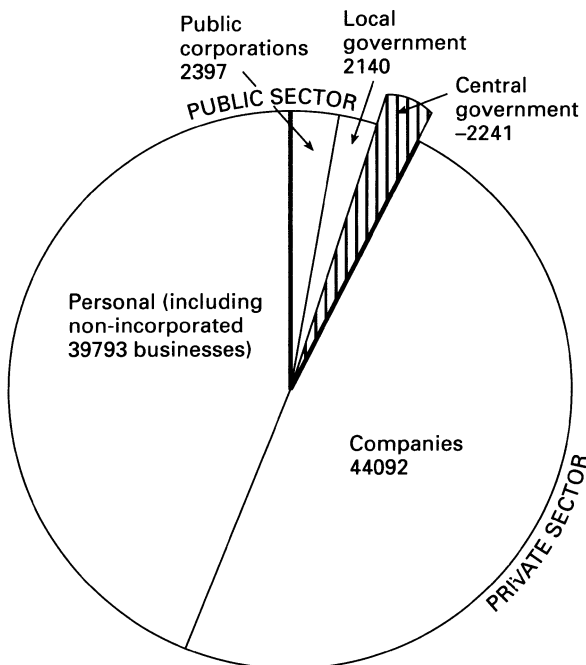
Figure 18.4 Saving in the UK, 1991 (£ million)

(b) Business saving

Saving by businesses (which in volume remains fairly stable) is achieved by not distributing all the profits made in a year. Some profits are usually retained, either to be 'ploughed back' for the expansion of the business, or to be held as liquid reserves in order to meet tax liabilities or maintain dividends when profits fluctuate. The chief factors affecting this type of saving are:

(i) *Profits*

Transfers to reserves are dependent upon and stimulated by the level of current profits. In practice, therefore, business saving is determined principally by the level of AD.

(ii) *Subjective factors*

Profits are likely to be retained when directors are expansion-minded or financially prudent.

(iii) *Government policy*

An increased tax on *distributed* profits would tend to increase company saving.

(c) Government saving

Central-government saving is achieved chiefly through a 'budget surplus', revenue exceeding current government expenditure. The surplus may be necessary: (i) to provide for the government's own investment and loans; and (ii) to ensure that, with personal and business saving, total saving will be sufficient to prevent an inflationary AD from developing (see Chapter 20).

Recently central-government expenditure has been so high that the difference has had to be covered by borrowing. A Public Sector Borrowing Requirement (see p. 266) means that there is dis-saving. In 1991 this amounted to £2241 m.

Public corporations are similar in many ways to ordinary businesses. But, as their operations are more directly under government control and their capital requirements are largely covered by the Treasury, their saving and investment are included under the public sector.

Local authorities, too, may have a budget surplus. In 1991 this was £2140 m.

Thus in the public sector saving is determined chiefly by government policy, economic and political.

(d) Conclusion

In the private sector, spending (and therefore saving) depend upon (i) the level of Y, i.e. the size of AD, and (ii) other factors influencing the amount spent out of income. In comparison with changes in AD these other factors are fairly stable. Hence the main factor affecting short-term changes in consumption spending is the size of AD!

We have therefore to look elsewhere for the reason why AD changes. It is to be found in the comparative instability of the other form of spending – investment.

18.4 Investment spending

(a) What do we mean by 'investment'?

Investment, for purposes of national-income calculations, is spending over a given period on the production of capital goods (houses, factories, machinery, etc.) or on net additions to stocks (raw materials, consumer goods in shops, etc.).

It is important to distinguish between this definition of *real* investment and *asset* investment – putting money in a bank or buying securities. In national-income analysis, investment takes place only when there is an actual net addition to capital goods or stocks.

Investment takes place in both the private and public sectors.

(b) Investment in the private sector

A firm will only spend on investment if it thinks that it will eventually prove profitable. There are, therefore, two main considerations to bear in mind: (i) the expected yield from the investment; and (ii) its cost.

The *yield* on the investment will depend largely on the demand for the consumer good the firm produces (see p. 149). In *estimating* such demand, the firm is most likely to commence from a position regarding which it does have some definite knowledge, i.e. the present demand for those goods. If that demand is buoyant and has remained so for a fairly long period, the firm will probably view the future optimistically. On the other hand, if present demand is low and has shown itself resistant to attempts to increase it, the future, to say the least, will appear somewhat gloomy. But since the current demand for goods, i.e. the level of consumption, depends chiefly on the level of income, we can say that, the higher AD is, the greater investment is likely to be.

Technical developments, like the internal-combustion engine, atomic energy, automation, computer technology and North Sea gas and oil, give an added impetus to investment.

Furthermore, the effect of government policy has to be considered. Changes in policy add to uncertainty. Is corporation tax likely to be increased? Will inflation compel the government to carry out restrictive policies? In contrast the government may stimulate private investment by subsidies or generous tax allowances, and revive optimism by increasing its own investment.

This brings us to the *cost* of investment. This is represented by the cost of borrowing money to finance it – the rate of interest. A low rate tends to stimulate investment. If the rate rises, marginal projects cease to be profitable, and so the level of investment falls (see p. 149).

However, whether the rate of interest has a *major* influence on investment is doubtful. For one thing, investment decisions, especially for large firms, are the result of long-term planning. Any alteration of plans just because of a change in the rate of interest might throw the whole programme out of phase. For another, firms allow a considerable safety margin when deciding on investment, probably expecting to recover its cost within five years. This margin is thus sufficient to absorb a relatively small rise in the rate of interest. Even the holding of stocks may not be affected by the rate of interest. Convenience is more likely to decide

the minimum held. In any case the rate of interest may be only a small part of the cost of holding stocks – warehousing, etc., being relatively far more important. Above all, compared with firms' expectations, the rate of interest is of secondary importance. Thus, especially when it comes to reviving investment, a fall in the rate of interest may have its main impact through the psychological improvement in expectations.

(c) Investment in the public sector
Capital expenditure is incurred by the central government, nationalised industries and local authorities.
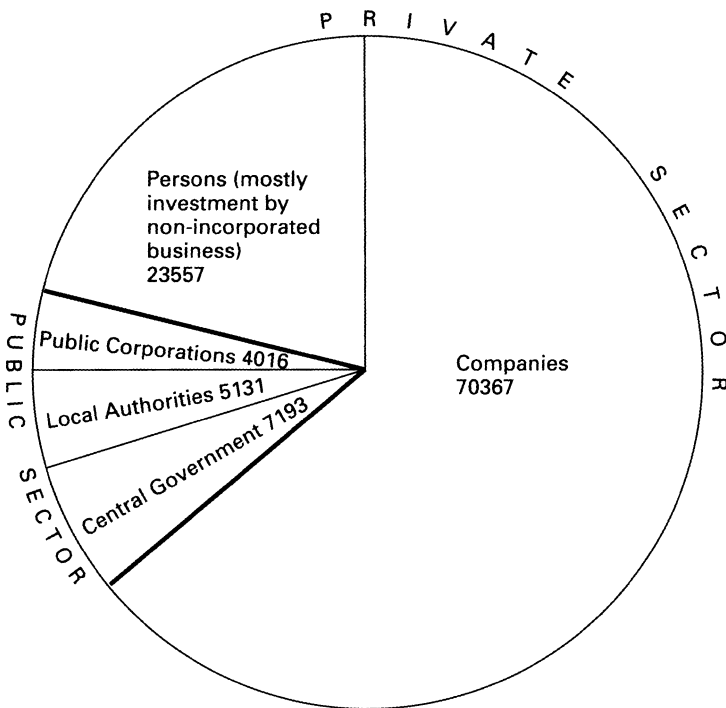
Figure 18.5 Investment in the UK 1991 (£ million) (gross domestic fixed capital formation and the increase in value of stocks and work in progress)

Much of central-government investment is fairly stable, depending chiefly on policy commitments. To a large extent, too, this is also true of the nationalised industries.

Local-authority investment tends, however, to react to changes in the rate of interest. This is especially true of spending on new dwellings. If, after applying

government grants, the cost of borrowing is not covered by the rents charged, local taxes have to be raised with adverse political consequences.

The real importance of public investment is that it is subject to direct government control. Thus, should private investment be deficient, the government can increase its spending on its own capital projects.

(d) Summary

Employment depends upon the level of AD – the total amount of money spent on the goods produced. AD fluctuates according to the relationship between intended saving and investment.

  (i) AD expands if:
  (1) investment increases but saving remains unchanged;
  (2) saving decreases but investment remains unchanged.
 (ii) AD contracts if:
  (1) investment decreases but saving remains unchanged;
  (2) saving increases but investment remains unchanged.

In practice investment is more liable to frequent change than is saving. Whereas firms' expectations are highly sensitive to new conditions, people's spending habits are fairly stable. Fluctuations in the level of AD, and therefore of income, are thus mainly the result of changes in the level of investment.

There is another important way in which saving differs from investment in the process of income creation. Whereas an increase in investment will, other things being equal, automatically produce an increase in saving through an expansion of income, an addition to saving need not lead to an increase in investment. Instead income merely contracts until what is saved from it equals investment.

18.5   The effect on the level of income of changes in investment

(a) The 'multiplier'

Given no change in saving, an increase in investment spending will raise the level of income by more than the increase in investment. Firms producing capital goods will take on additional workers. A part of these workers' earnings will be spent, providing additional income for shopkeepers and others, who in turn will spend a part of their extra income. There is thus a 'multiplier' effect. How the process works in real life can be illustrated from Nevil Shute's *Ruined City*. The shipyard in the town obtained an order for three tankers. 'A shop, long closed, reopened to sell meat pies . . . A man who gleaned a sack of holly in the country lanes disposed of it within an hour . . . A hot roast chestnut barrow came upon the streets, and did good trade.'

The multiplier is defined as the

$$\frac{\text{increase in income}}{\text{increase in investment}}$$

Its size depends upon the fraction of additional income which is spent, that is, the marginal propensity to consume (see p. 228). Suppose investment increases by

£100 million a year, and that households spend half of any additional income. Producers of capital goods can now pay an extra £100 million on wages, rent, interest and profit. Households receiving this extra income spend £50 million with shopkeepers who, in turn, spend £25 million. The sum of these additional incomes will eventually total £200 million. The multiplier equals 2.

To generalise, the multiplier equals

$$\frac{1}{1- \text{the marginal propensity to consume}}$$

If, in the above example, the marginal propensity to consume had been $^3/_4$ instead of $^1/_2$, income would have increased by £400 million, giving a multiplier of 4. Since the marginal propensity to consume + the marginal propensity to save = 1, we can write the multiplier as equalling $1/s$, where s equals the marginal propensity to save.

(b) Diagrammatic representation

Figure 18.6 presents the essentials of the determination of the level of income. It is assumed that investment is autonomous, that is, its level is independent of the level of income (as shown by the horizontal line I). Saving varies with income. Initially saving is negative – households maintain spending at a low level of income by drawing on accumulated savings or by borrowing. When income increases and dis-saving becomes unnecessary, only $^2/_3$ of additional income is spent, $^1/_3$ being saved as shown by the line S having a constant slope of $^1/_3$.

The equilibrium level of income is Y, where $S=I$. At any income less than this, AD will expand because I is greater than S; and vice versa.

If now investment increases to I', there will be a new equilibrium level of income, Y'. Diagrammatically, the multiplier equals YY'/II', i.e. 3. Similarly, income would decrease if the level of investment fell below I.
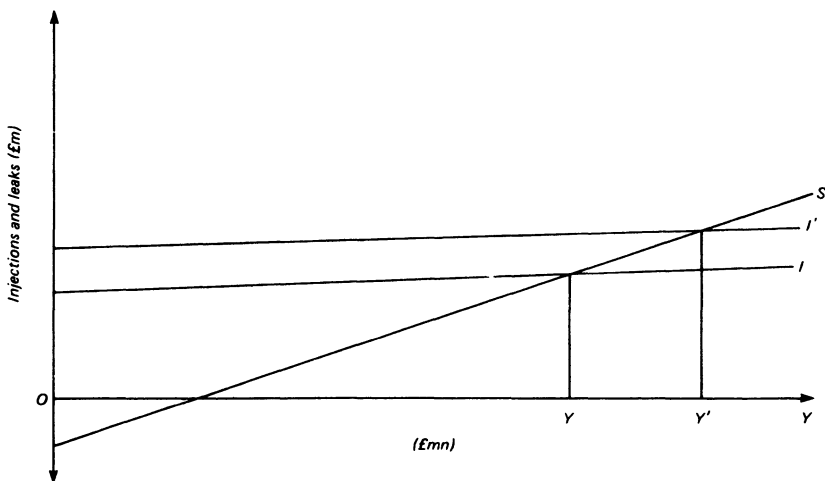
Figure 18.6 The effect on income of an increase in the level of investment

18.6   Leaks and injections in general

In the above discussion, saving has been the only leak from the flow of aggregate demand. Similarly investment has represented the sole injection. However, the relaxation of the assumptions that there is no government and a closed economy introduces the consideration of other leaks and injections.

Direct taxation takes income from households, leaving them with less available spending power. Indirect taxes siphon off a part of consumer spending to the government, thereby reducing firms' receipts. Thus taxation is a leak from the circular flow of income. On the other hand, it can be put back by government spending – an injection.

Finally we have to consider the effect of imports and exports. When British households spend on foreign goods and services, there is a leak of AD from the circular flow, since this spending is received by foreign firms. On the other hand, spending by foreigners on British goods and services is received by British firms, and therefore represents an injection (Figure 18.7).

These other leaks – taxation and imports – and injections – government spending and exports – affect AD and Y in the same way as do, respectively, S and I.
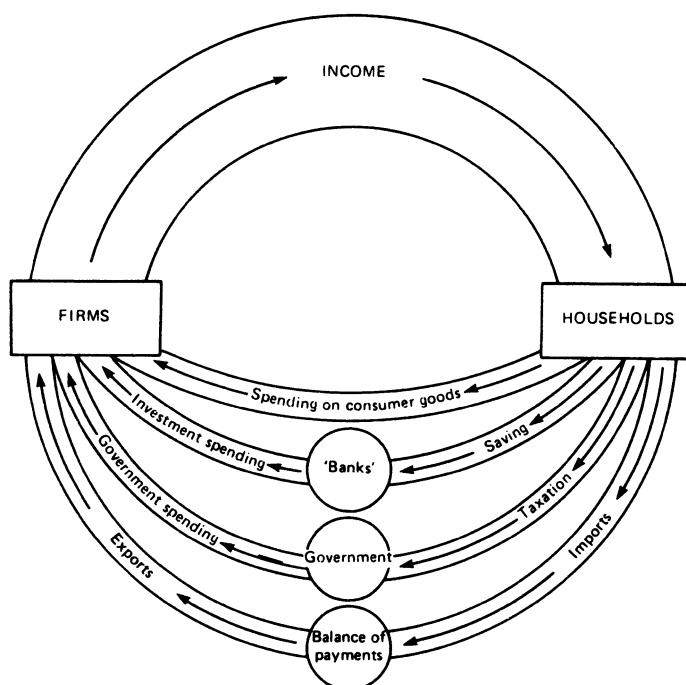
Figure 18.7   Total leaks and injections

18.7   Aggregate demand and full employment policy

The above is an outline of Lord Keynes' theory of how the level of employment is determined (*The General Theory of Employment, Interest and Money*, 1936).

Employment depends upon the level of national income (Y). If the level of AD is too low, the economy will be in equilibrium where Y is below the full-employment level (Figure 18.8). Thus, Keynes considered, the government's task is to estimate that level of AD which will produce full employment, and then arrange that AD is increased to this level.
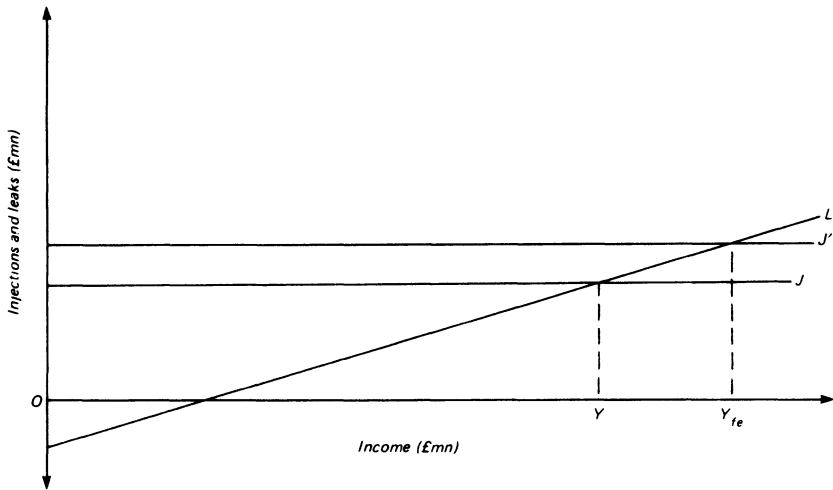
Figure 18.8   Achieving a full-employment aggregate demand by budgetary policy

This could be helped by an appropriate monetary policy, given that a lower rate of interest would induce investment, and easier credit terms stimulate consumer spending.

But the main and more effective weapons were fiscal and budgetary policies.

Fiscal policy can influence private consumption and investment indirectly by simply changing the *type* of taxes levied. Thus a switch from indirect taxation would tend to increase consumption, for it would mean greater spending power for poorer people (those having a high propensity to consume). Similarly reduced taxes on businesses would tend to increase investment through improved profitability.

More directly, AD may be influenced by *budgetary policy* which adjusts the relationship between government taxation and expenditure. As we have seen, taxation represents an appropriation by the government of a part of private incomes, and will be retained in the circular flow of income only in so far as it is spent by the government. Hence AD will be increased if taxation is less than

government spending, and vice versa. If previously the budget was balanced, there would now be a budget deficit, and vice versa. In Figure 18.8, current Y is OY, less than the full-employment OY_{fe}. It is therefore necessary to increase total injections from J to J', e.g. by increased government spending or additional private consumption.

Thus today the budget is regarded, not simply as the means of raising revenue to meet the year's estimated expenditure, but as a weapon to adjust private spending power to the output which available resources can produce given the demands of the public sector. Since the latter are largely determined by policy commitments, the ability to vary public expenditure is limited, and changes in taxation, therefore, are the main means of 'fine-tuning' the economy. Reducing direct taxation increases disposable income and, provided this increase is not entirely saved, spending will increase (expanding income according to the multiplier).

18.8   Defects of demand management

(a) Post-war demand management in the UK
Keynes's theory dominated economic policy during the 1950s and 1960s. Successive UK governments maintained an AD which until 1974 (apart from 1971–2) kept unemployment below 3 per cent. Whenever expansion of AD led to balance-of-payments difficulties, the government imposed temporary 'stop' policies.

But 'fine-tuning' the economy to achieve a high level of demand produced a state of 'overfull' employment, with vacancies consistently exceeding unemployed workers. The outcome was inflation, balance-of-payments difficulties, shortages of skilled workers, high labour turnover, firms holding on to surplus workers in case they should be required in the future ('disguised' unemployment), underinvestment in new equipment, a failure by industry to adopt new techniques or to switch to the production of high-technology products and an unsatisfactory rate of growth (see Chapter 26). As a result, British industry was in poor shape to withstand international competition during the world recessions of 1974–5 and 1979–82.

More than that, in the 1970s the rate of inflation and the level of unemployment increased together – a situation known as 'stagflation'. Hence Keynes's views came under increasing scrutiny.

(b) Weaknesses of the Keynesian Approach
It must be remembered that Keynes was writing his *General Theory* against the background of the high cyclical unemployment of the 1930s. For him, therefore, unemployment was the major problem. In reality, though, the many objectives of government stabilisation policy are interrelated and cannot be dropped into different boxes with separate measures to deal with each.

First, too little attention was given to how supply responds to increases in AD as full employment is approached. The analysis assumes a stable price level. This is justifiable at higher rates of unemployment. Firms have spare capacity and can

increase output at constant costs. Increased AD is therefore covered by increased output and the price level remains stable. But eventually, as output increases, firms experience rising costs owing to more intensive use of existing capacity and bottlenecks in obtaining variable factors. As a result prices rise (see p. 245).

Keynes recognised that this rise in prices would occur: the remedy was to remove the inflationary pressure by reducing AD. But this ignores the dynamic forces which come into play once prices start to rise, chiefly the role of trade union inflationary expectations in the wage-bargaining process (see Chapter 19).

Second, Keynes's theory of the price level failed to recognise that increases in the money supply could have the effect of increasing AD directly (see pp. 248).

Third, Keynes underplayed the side effects of a large PSBR which high government spending could give rise to. Keynes recognised that government deficit-spending would increase the National Debt, but since loans could be 'rolled over' as they matured, only interest had to be currently paid and, apart from overseas lenders, such payments were simply transfers from taxpayers to lenders within the UK. But Keynes did not foresee the big increase in government borrowing necessary to pay for the welfare state. Servicing a large PSBR creates difficulties for both monetary and fiscal policy. If it is covered by borrowing from the banks, their extra liquidity allows them to increase the supply of money. If, alternatively, it is covered by non-bank borrowing, e.g. by institutions or persons buying government bonds, the price has to be sufficiently attractive to clear the market. Thus the rate of interest rises. This could cause private investment to fall – 'crowded out' by the increased government spending!

Fourth, changes in the level of AD and in the price level have repercussions on the balance-of-payments which may impose a constraint on the expansion of AD.

The government's difficulties in devising policies which seek to achieve greater harmony between its objectives will be developed in later chapters. Our first task is to consider why inflation should be controlled (Chapter 19). We can then develop a model which relates the level of activity to the price level, explains the inflationary process, and suggests policies which could help in achieving a lower rate of unemployment consistent with price stability.

19 The effects of inflation

19.1 Why control inflation?

Inflation can be defined as a sustained rise in money prices generally. Today the control of inflation is given priority in government policy. To appreciate why, we have to look at the effects of rising prices or – what is the same thing – a fall in the value of money.

(a) Possible benefits

At one time a gently rising price level was not viewed with too much concern. It improved the climate for investment and so helped to maintain AD. Moreover, it tended to reduce the real burden of servicing the national debt: while interest is fixed in money terms, receipts from taxation increase as national money income rises.

The snag, however, is that, once started, the rise in prices is difficult to contain. At first it becomes uncomfortable, producing undesirable results both internal and external. Eventually, the rate of inflation increases. The situation is then serious, for it is much more difficult to reverse the trend. Indeed, it can develop into runaway inflation.

(b) Internal disadvantages

   (i) Real income and wealth are re-distributed arbitrarily. Not only does inflation reduce the standard of living of those dependent on fixed incomes, e.g. pensions not 'inflation-proofed', but it benefits debtors and penalises lenders. Thus the stability upon which all lending and borrowing depends is undermined.

  (ii) Interest rates rise, both because people require a higher reward for lending money which is falling in value and also because the government is forced to take disinflationary measures. Not only may this discourage investment but it can have social consequences, e.g. in meeting monthly mortgage payments.

 (iii) Saving is discouraged because postponing consumption simply means that goods cost more later.

 (iii) The allocation of resources is distorted. Thus institutions, such as insurance companies, invest funds in assets having a strong inflation hedge. As a result the capital value of such assets rises, encouraging developers, for instance to build office blocks rather than houses for letting.

 (iv) Efficiency is reduced because:

    (1) a buoyant sellers' market blunts competition as higher selling prices allow even inefficient firms to survive;

    (2) uncertainty is increased;

    (3) market signals are less clear since some inflation-hedge may be included e.g. in contracts;

    (4) it strengthens the possibility of disruption of production until demands for wage increases are agreed;

    (5) financial services spring up to advise on protecting savings from losses through inflation.

  (v) Inflation generates industrial and social unrest since there is competition for higher incomes. Thus, because of rising prices, trade unions ask for annual wage rises. Often demands exceed the rate of inflation, anticipating future rises or seeking a larger share of the national cake to improve their members' real standard of living. Those with the most 'muscle' gain at the expense of weaker groups.

  (vi) Additional administrative costs are incurred in off-setting go-slow and work to rule disruptions, allowing for inflation in negotiating contracts and wage rates, revising price lists and labels, etc.

  (vii) The rate of inflation tends to increase, largely because high wage settlements in anticipation of higher future prices help to bring about the very rise which people fear.

(c) External effects

Inflation can create balance-of-payments difficulties for a country dependent on international trade, as Britain has discovered over the past forty years.

  (i) Exports tend to decline because they are relatively dearer in foreign markets. Firms exporting a large proportion of their output may therefore have to lay off workers.

  (ii) Imports tend to increase because foreign goods are relatively cheaper on the British market.

  (iii) Higher money incomes in the UK increase the demand for imports and tend to decrease exports because the buoyant home market makes it less vital for manufacturers to seek outlets abroad for their goods.

  (iv) An outward movement of capital may take place if price rises continue since foreign traders and financiers lose confidence in the pound sterling maintaining its current rate of exchange.

While the above effects are uncomfortable, it is possible to live with a moderate rise in prices. The snag is that where rising prices are thought likely to continue, people bring forward their spending, thereby producing the very price rise feared – an example of 'self-justified expectations'. So the process gathers momentum, stimulated still further by demands for higher wage increases.

19.2   Measuring changes in the general level of prices

(a) Difficulties

The difficulty of measuring changes in the general level of prices is that different kinds of price – wholesale prices, retail prices, security prices, import prices, etc. – change differently. If we tried to measure changes in all prices, therefore, our

task would be stupendous. More than that, it would lack practical significance; a measurement which, for instance, included security prices would have little interest to a working man owning no securities. It is usual, therefore, to concentrate on changes in the prices of those goods which are of most general significance – the goods bought by the majority of people.

(b) Method of measuring changes in the value of money

Since we are mainly interested in the extent to which the value of money has altered between one date and another, it can be measured as a relative change by means of an *index number*. The steps are as follows:

(i) A base year is selected. This is now referred to as the 'reference date'.

(ii) In order to ensure that the same goods are valued over the period under consideration, a 'basket' of goods, based on the current spending habits and income of the 'typical' family is chosen.

(iii) The basket is valued at reference date prices, and expressed as 100.

(iv) The same basket is revalued at current prices.

(v) The cost of the current basket is then expressed as a percentage of the base year. Thus if the cost of living had risen by 5 per cent, the index for the current year would be 105.

In practice, the percentage price relatives of the selected goods are calculated, and are then 'weighted' according to the relative expenditure on the commodity at the reference date. Suppose, for instance, that there are only two commodities, bread and meat, upon which income is spent. The index between two years is calculated as follows. The price in year II is expressed as a percentage of the price in year I. This is multiplied by the appropriate weight to give a 'weighted price relative'. These weighted price relatives are then totalled and divided by the total of the weights to give the new index number.

	YEAR I				YEAR II Price relatives		
	Price	Units bought	Expenditure	Weight	Price	Year II as % of Year I	Weighted price relative
Bread	30p	5	150p	10	45p	150	1 500
Meat	150p	11	1 650p	110	180p	120	13 200
				120			120) 14 700
							122.5

Index
Year I (base)   100
Year II   122.5

Difficulties in calculating index numbers

The method outlined above of calculating changes in the value of money has obvious snags:

(1) The basket and the weighting are merely an arbitrary average. Different income groups have widely different baskets, and even within the same group the

amount spent on each good varies. Thus a change in the Retail Prices Index does not affect all people equally.

(2) The basket becomes more unreal the further we move from the reference date (at present 13 January 1987 = 100). For instance, an increase in income gives a different pattern of expenditure, new goods are produced and the quality of goods changes, and spending is varied according to relative price changes. The Retail Prices Index tries to overcome this defect by revising the weights each January on the basis of the *Family Expenditure Survey* for the previous year.

(3) Technical difficulties may arise both in choosing the reference date and in collecting information. For instance, the reference date may coincide with abnormally high prices, while the development of discount stores may upset standardised methods of collecting prices.

Thus a Retail Prices Index is merely an indication of changes in the cost of living. But if we bear its limitations in mind, it is the most useful measurement we have of changes in the value of money. It therefore provides the yardstick for calculating changes in real earnings and for 'inflation-proofing' public sector pay and pensions.

20 Employment, price stability and government policy

20.1   Early views on the causes of inflation

(a) Aggregate supply and aggregate demand curves

(i) *The Aggregate Supply curve (AS)*
The AS curve shows the relationship between the price level and the real output
of the economy. It is, however, a short-run supply curve in the sense that produc-
tive capacity is fixed: output can be increased only by increasing inputs of the
variable factors, chiefly labour.

There are three different situations:

1. At a low level of output, supply can be increased at a constant cost. Firms
   have surplus capacity and equally efficient labour, basic materials, com-
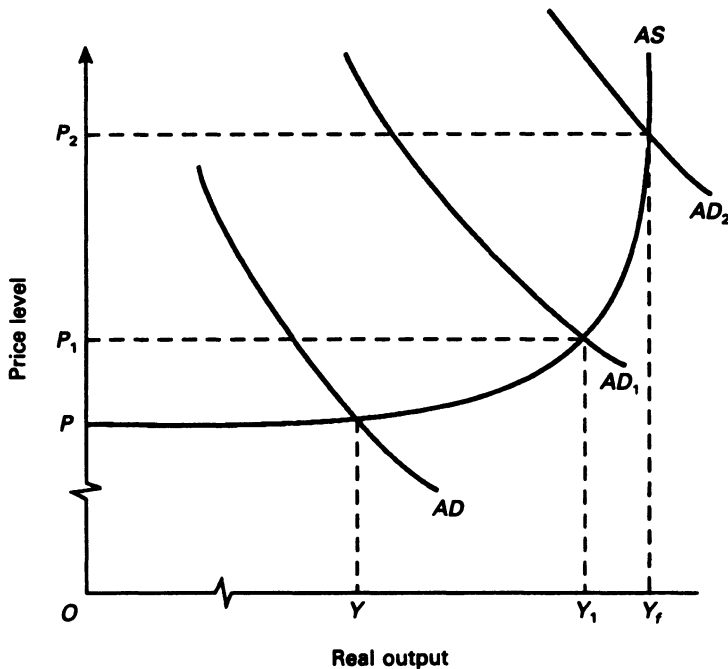
Figure 20.1 The relationship of output and the price level to AD and AS

ponents, etc. can be obtained at existing prices. Thus the AS curve is perfectly elastic up to output *OY* with the price level at *P* (Figure 20.1).

2. As output increases beyond *OY*, fixed capital has to be used more intensively and diminishing returns set in. In addition 'bottlenecks' appear through skilled labour shortages, longer delivery dates for components, etc. Above all, trade unions are in a stronger position and negotiate wage increases.

3. The AS curve slopes upwards at an increasing rate the nearer output approaches OY_f – the maximum potential output. Given this absolutely inelastic AS curve at OY_f any attempt to expand output by increasing AD will be impossible and will simply find its outlet in steeply rising prices.

(ii) *The Aggregate Demand curve (AD)*

As we have seen (Chapter 18), AD consists of consumption, investment and government spending, together with exports less imports – in symbol form: $C + I + G + (X - M)$. In the presentation here of this new analytical approach, however, increases or decreases in AD have to be shown by *shifts* of the curve. Nevertheless we still have to explain the *shape* of the curve, that is, how AD is related to the price level. To simplify we will concentrate on *C* and *I*.

The AD curve slopes downwards from left to right (Figure 20.1) showing that, as the price level falls, AD expands. In other words, the lower the price level, the greater will be the total real output demanded in the economy.

The reasons for this are:

1. *Consumers feel wealthier* since balances held in cash will buy more (the 'real balance effect'). As a result spending is likely to increase.

2. *The rate of interest falls.* A fall in prices means that less money is required for everyday spending. With no change in the money supply, more money is left for buying, not only goods, but also financial assets such as government bonds. The resulting rise in bond prices represents a fall in the rate of interest (see p. 314).

3. *The balance of payments* $(X - M)$ *improves* as home-produced goods become more competitive, increasing exports and decreasing imported substitutes.

(iii) *AD and AS combined*

To simplify, we will assume no government and a closed economy.

Equilibrium occurs at the intersection of the AD and AS curves because at this point the output which households and firms are willing to spend on *C* and *I* equals the output which firms will supply.

(b) Causes of inflation: a first approach

Prices rise when there is excess purchasing power for goods available at current prices. But what brings about the excess purchasing power? Early explanations were often divided into *demand-pull* and *cost-push*.

(i) *Demand-pull*

When an increase in AD causes output to expand above *OY* (Figure 20.1), prices start to rise. Thereafter, if AD continues to increase, the price rise accelerates.

Indeed should AD increase after the full employment output (OY_f) has been reached, we have true inflation, the extra AD finding its outlet entirely in higher prices.

(ii) *Cost-push*
The price rise can start on the supply side through a rise in import prices. More usually, however, it follows from demand pull as labour seeks compensating wage increases. Indeed wage demands may exceed the current rate of inflation even though there has been no corresponding increase in productivity. Employers tend to concede these wage demands, and raise prices accordingly. Such price rises can be absorbed by the higher earnings. Here, therefore, prices have been pushed up by an initial increase in costs.

(c) Policy implications
The foregoing simple explanation of the direct relationship between output and rising prices seemed to be supported in 1958 by Professor A.W. Phillips whose research indicated a strong negative relationship between the annual rate of inflation and the annual rate of unemployment in the UK over the past century (Figure 20.2).
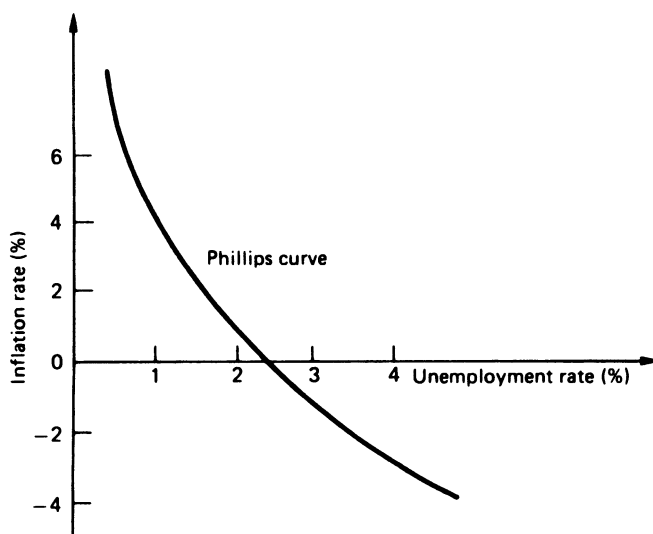
Figure 20.2   The Phillips curve

The policy conclusion was that a reduction in the rate of inflation could be 'traded off' by reducing AD though this could lead to higher unemployment. Indeed this was the essence of the 'stop-go' policy followed by the UK in the 1950s and 1960s whenever rising prices led to balance of payments problems.

Alternatively, where the basic cause was cost-push, the government urged wage restraint, backing this with restrictions on price rises.

(d) Weaknesses of the above explanation of inflation

Until the mid-1960s the Phillips curve relationship held, but then both unemployment and inflation increased together, a situation described as 'stagflation'.

The weaknesses of the above analysis is that it suffers from being too static. As a result it sees inflation as being a *condition* whereas it is really a *process*.

First, unlike the demand and supply curves which we studied in Chapter 3, the AD and AS curves are not independent of each other. Thus in Figure 20.3 an increase in AD to AD_1 produces a rise in the price level. As a result, increased wage rates are negotiated, so that the AS curve moves to AS_1. But since this also means an increase in money spending power, aggregate demand moves from AD_1 to AD_2 at a higher price level. This, by generating further wage catching up, produces an inflationary spiral.
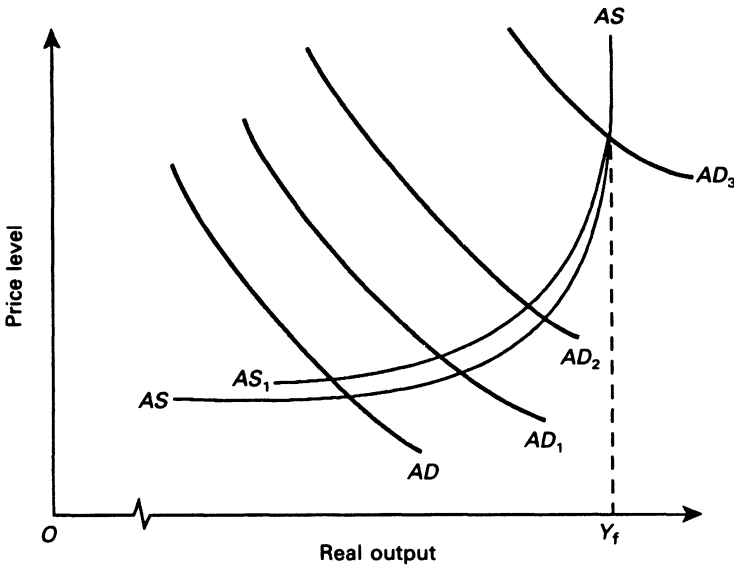
Figure 20.3 The inflationary process

Second, wage demands may eventually be geared to an *expected* higher rate of inflation. The rate of the inflationary spiral now gathers momentum. Thus the remedy is not a simple piece of surgery to remove excess fat, but rather a fight against a cancerous growth.

The experience of 'stagflation' showed that an alternative theory of inflation embodying expectations was required. This has been built up by the 'monetarists' led by Professor Milton Friedman.

20.2   Monetarism

The monetarist theory of inflation emphasises: (a) the *direct impact* of the money supply on AD; (b) *expectations* as the cause of the inflationary spiral; and (c) *the natural rate of unemployment*. We will explain each in turn.

(a) The money supply and AD

It was observed that there was a positive correlation between increases in the money supply and the rate of inflation. But is there a causal connection? Does an increase in the money supply *directly* increase AD?

Keynes considered that the supply of money does not enter directly into spending decisions. In the short run these depend on the level of income; over time they will be affected by long-term factors, such as contractual commitments to regular saving. Any increase in the supply of money simply increases liquidity in the economy and the rate of interest – the price paid for liquidity – therefore falls. Aggregate demand (AD) will expand only indirectly, through a lower rate of interest leading to more investment spending.

In contrast, Milton Friedman holds that an increase in the supply of money can lead *directly* to additional spending and thus *cause* inflation. He argues that people maintain a fraction of their nominal income in cash balances. An increase in the money supply results in their having larger cash balances than they require, and so run them down by spending. Such spending increases AD and money incomes until cash balances are equal to their former fraction. Nor does this surplus cash have to be spent on 'bonds', Keynes' omnibus term for non-money assets. Wealth can be held in many forms: cash which yields liquidity, 'bonds' which yield interest and possible capital appreciation, and, the Monetarists emphasise, *consumer goods* which yield utility. People distribute their spending according to their marginal preferences for these different forms of yield (which in turn can be influenced by their expected rate of inflation).

Thus any increase in the money supply is likely after a little while to lead to some increase in the demand for consumer goods, resulting in a rise in their prices.

(b) Inflation expectations

Once inflation expectations enter into wage negotiations, the monetarists hold that increasing AD will not achieve a *long-term* decrease in the rate of unemployment but simply result in higher inflation. In short there is no trade-off between inflation and unemployment.

Suppose that the current rate of inflation is 4 per cent and that the rate of unemployment is 5 per cent. This is a stable position because wage-bargainers have expected 4 per cent to be the inflation rate.

The government now decides that it wants to reduce the unemployment rate to 4 per cent and accordingly increases AD. Prices rise to, say, 6 per cent inflation, but money wage rates do not rise. Thus firms, enjoying increased profitability, increase output so that initially unemployment falls to the desired 4 per cent.

But this is only a short-term position depending on the fact that workers tend to concentrate on nominal money-wages rather than real wages – the 'money

illusion'. Eventually, however, they realise that real wages have fallen and in their next wage negotiations they obtain at 6 per cent rise in money wages to cover what they expect to be the rate of inflation. The recovery in real wage-rates increases costs, so that firms reduce output and unemployment reverts to 5 per cent – but the rate of inflation has risen to 6 per cent.

(c) The 'natural rate of unemployment'

This 5 per cent (in this case) the monetarists would term the 'natural rate of unemployment' – the rate to which unemployment will tend given the current wage rate. If this wage rate is above that which would clear the labour market when there is equilibrium *throughout the economy*, the supply of labour exceeds the demand; that is, there are workers offering their services at the current wage rate who cannot find a job. The natural rate of unemployment refers to this involuntary unemployment.

(d) Policy implications of monetarist theory

Monetarist theory originated as an explanation of how increases in the supply of money led directly to increases in AD and thus to rising prices. Even with cost-push inflation, extra money is necessary to accommodate wage increases. Thus the first condition for keeping inflation in check is strict control over the money supply.

A major difficulty here is that 'money' is difficult to define (see p. 175) and policy has to be based on an imperfect money index (see p. 313). But, as we have seen, the theory goes further to show that expanding AD to reduce unemployment will only be successful for as long as there is a fall in real wage rates. This may occur in the short period owing to the 'money illusion', but eventually labour incorporates a revised higher expected rate of inflation in its wage negotiations.

Government policy must also be directed to convincing the trade unions that the trend of the future rate of inflation is downwards. This is not easy for until wage negotiators can actually see a falling rate of inflation, they interpret a wage increase which falls short of their expected rate of inflation as a cut in real wages. Yet voluntary wage-restraint is the 'least-cost' policy for controlling inflation. The difficulty is that voluntary restraint soon breaks down.

An alternative approach is to reduce inflation expectations by announcing strict monetary and fiscal targets for the medium-term future. The extent to which such a policy reduces inflation expectations largely depends on how far unions are convinced of the government's resolve to hold to its targets. In this respect it should be noted that the government can indicate its determination by the extent to which it restricts wage increases in the public sector.

Unfortunately the change in the economic climate takes time to be realised especially when in the past trade union negotiators have been consistently successful in obtaining inflationary wage increases. In the meantime there may be a sequence of unsuccessful strikes and a prolonged period of unemployment. This is the 'high cost' alternative.

How actual policies have been implemented in the UK in recent years are described in Chapter 26.

20.3   Supply-side measures

The introduction of the hypothesis of a 'natural rate of unemployment' suggests that, since expanding AD is ineffective in reducing unemployment, less emphasis should be placed on demand and more attention given to supply. Reducing the costs of production would mean that a greater output could be supplied at any given price level.

The strategy is illustrated in simplified form in Figure 20.4, which concentrates on the rising portion of the short-run aggregate supply curve.
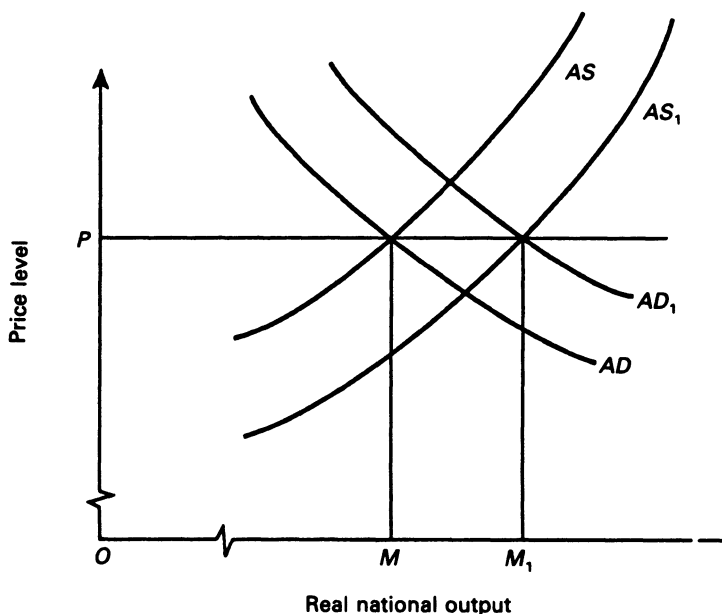
Figure 20.4 Supply-side policy

If, for example, the curve can be shifted from AS to AS_1, aggregate demand can be increased from AD to AD_1 so that output expands from OM to OM_1 with no rise in the price level.

Policy to achieve this embraces a variety of incentives usually included under the umbrella term 'supply-side economics'. It integrates longer-term micro measures within the overall macro policy. They seek to achieve greater market freedom and efficiency, a reduction in costs, and incentives to increase effort, reward enterprise and encourage investment.

(a) Market freedom may be enhanced by:
   (i)   freeing them from government controls (e.g. incomes policy, minimum wage regulations, pricing policies of the nationalised industries);
  (ii) promoting competition (e.g. in real property conveyancing, optical and financial services);

(iii) restricting the power of trade unions (e.g. through precipitate and sympathetic strike action, unofficial strikes, the closed shop);

(iv) the privatisation programme;

(v) introducing competition in the natural monopolies by new devices;

(vi) removing institutional barriers in the capital market (e.g. exchange control, the Stock Exchange);

(vii) using the rate of interest (the price of liquid capital) as the main weapon for adjusting aggregate demand.

(b) Costs reduction include:

(i) lowering national insurance contributions (an *ad valorem* tax on employing labour);

(ii) improving the supply and quality of labour, e.g. by developing training schemes, encouraging flexible working hours to attract part-time workers, improving mobility, granting subsidies to firms to locate in the Assisted Areas;

(iii) providing advisory services (e.g. on the EC's Single Market).

(c) Incentives cover:

(i) a reduction in the marginal rate of income tax on effort, etc. even though it may reduce total tax revenue;

(ii) a lower corporation tax to encourage investment and the taking of entrepreneurial risks;

(iii) special help for new firms to obtain the initial capital, e.g. 'start-up' schemes, the encouragement of 'venture capital' institutions;

(iv) profit-related pay (which gives employees a direct stake in the success of the company and enables pay to respond more readily to changing market conditions), share option schemes and wider share ownership generally.

But although supply-side measures are necessary, it has to be recognised that they take time to be fully effective and are unlikely to be sufficient. They must not be regarded as an alternative to the over-riding objective of obtaining from labour a working agreement on moderating wage demands.

㉑ Balanced regional development

21.1 The nature of the problem

(a) Causes of regional depression

In broad terms, a regional problem can arise because:

1. *The particular region may be endowed with poor natural resources*
This applies, for example, to the Highlands of Scotland. More generally, with the growth of national income, an agricultural region which does not attract expanding industries – e.g. Cornwall and Devon – cannot provide its population with living standards comparable with those of the rest of the country.

2. *The resources of the region may not be fully developed, usually through lack of capital*
This applies particularly to the less developed countries. Here the more immediate solution is for capital to be provided on favourable terms by richer regions.

   In the long term an improvement in the imbalance may depend mainly on rising incomes elsewhere. Exploitation of the area's resources may now become economically viable. For example, prosperity came to Aberdeen and the Shetlands only when the rise in the price of oil and the development of modern technology made extraction of North Sea oil an economic proposition. Alternatively, rising incomes in other regions may allow tourism to be developed, e.g. North Wales, the Lake District.

3. *A region's basic industry is either stagnant or in decline*
Such a region is usually characterised by: a rising rate of unemployment; a level of income which is falling relatively to other regions; a low activity rate, particularly of female workers; a high rate of outward migration; and an inadequate infrastructure. It is thus this type of regional imbalance which creates the problem for *national* governments; indeed the depressed regions are normally identified by their unemployment rates (Table 21.1).

(b) Weaknesses of correction through the market mechanism

Theoretically the price system should move workers who become unemployed to other jobs. The fall in the demand for a good, and the consequent unemployment, should result in a relative wage fall. On the other hand, where demand is buoyant wages should rise. Such changes in relative wages should move: (a) workers from low-wage to high-wage industries, and (b) industries from high-wage to low-wage areas.

Table 21.1   Percentage rate of unemployment by region, July 1993

United Kingdom	10.4
Region:	
South East	10.2
East Anglia	8.4
South West	9.8
West Midlands	11.0
East Midlands	9.5
Yorks. and Humber	10.3
North West	10.8
North	12.2
Wales	10.3
Scotland	9.7
N. Ireland	14.0

Source: *Employment Gazzette* (Department of Employment).

The attraction of this model is that economic efficiency and the correction of an imbalance can be brought about by the market. Yet in suggesting that a government regional policy is largely superfluous, it has serious weaknesses:

(i)   Factor markets adjust much less perfectly than the theory implies. Labour is immobile (see pp. 136–8), while factor prices especially wage rates, tend to be resistant to any downward movement. Moreover, national wage-bargaining weakens the response by the price signals to regional imbalance.

(ii)   The assumption of constant returns to scale which is implicit in the theory may not hold. Manufacturing in particular is characterised by increasing returns to scale over the relevant output range so that high-wage regions may also generate high returns to capital. Thus firms, like labour, may migrate to the prosperous high-wage regions. Indeed, as communications improve, these regions may gain with the progressive opening of trade at the expense of the decaying region. Thus the south-east region of England has benefited from its close connections with the EC.

Should movement be entirely outwards, the model has additional weaknesses:

(iii)   The theory ignores the external costs to society of (i) the loss of social capital and the disintegration of communities in the depressed regions, and (ii) the congestion and inflationary pressures generated in the expanding areas (see below).

(iv)   Those workers who do move from the depressed regions are mainly the better educated, most highly skilled and more enterprising young adults. As such they are often the leaders of the community. The result is that the region becomes still further depressed *and* thus unattractive to new industries.

(v) The theory ignores the fact that migration from the depressed regions leads to a loss of income there. The multiplier effect of reduced consumer spending and investment serves to depress the area still further.

(c) Consequences of regional depression

The existence of prolonged depression in certain regions has adverse consequences which can be summarised as follows:

1. *An underutilisation of resources through unemployment*

Not only does regional unemployment result in lost output for the community as a whole, but it can also have serious social and psychological effects on the workers concerned. Moreover, significant differences in people's income between regions has equity implications.

2. *A loss of social capital as towns and cities decay*

Where the nation's population is static or falling, outward migration from depressed areas involves social costs in that schools, churches, etc. fall into decay while certain public services have to be operated below capacity. In contrast new roads and public buildings, such as hospitals, have to be provided in the expanding areas.

3. *External social costs*

Migration from decaying regions often results in a loss of welfare through the break-up of communities and the destruction of the 'social character' of an area.

Similarly, there may be external costs of excessive urbanisation (e.g. traffic congestion, noise, pollution and intensive housing) through migration to a prosperous region.

4. *Differences in unemployment between regions make it more difficult to manage the economy*

Prosperous regions tend to become 'overheated' through the pressure of demand. This is reflected in higher wage-rates and labour shortages. Higher wage-rates tend to be transmitted even to the depresses regions through national wage agreements, the insistence on traditional wage differentials, etc. But anti-inflationary measures, both monetary and fiscal, add to the unemployment problem of the depressed regions.

(d) Broad outlines of regional policy

The objectives of regional policy have widened in the light of experience. In brief, it now seeks to:

(i) reduce the relatively high level of unemployment in certain regions;
(ii) achieve a better balance between the population and the environment;
(ii) preserve regional cultures and identities;
(iv) relieve inflation by reducing the pressure of demand in the expanding regions;
(v) counter possible adverse regional effects of greater international economic integration and of more open economies.

Where an area is depressed, the government can give first aid by placing its contracts there e.g. for defence equipment, and awarding it priority for public-works programmes – schools, new roads, hospitals, the physical regeneration of urban areas, etc. Subsidies may also be granted to secure contracts, for example, to build ships.

In the long term, however, the government must take measures that will, on the one hand, encourage the outward movement of workers, and on the other induce firms to move in to employ those workers who find it difficult to move and also to halt further degeneration of the region. The first is usually referred to as 'workers to the work', the second as 'work to the workers'.

21.2 Workers to the work

Taking workers to the work is basically a micro approach to overcome market frictions, chiefly the immobility and imperfect knowledge of labour. In pursuing this policy, however, the government must bear in mind the following:

(a) Unemployment arising through immobility is far more difficult to cure when cyclical unemployment also exists, for an unemployed person has little incentive to move if there is unemployment even in the relatively prosperous areas.

(b) Other government interference in the economy may add to the problem of immobility. Thus high rates of income tax whittle away monetary induce-ments to move and unemployment benefit may reduce the incentive to seek a job elsewhere. Similarly, rent control and residential qualifications for local authority housing priorities lead to difficulties in finding accommodation.

(c) Even owner-occupiers in depressed regions may be restricted in mobility by the much higher cost of housing in the prosperous areas.

(d) Many changes of both occupation and area take place in a series of ripples. Thus an agricultural labourer may move to road construction to take the place of the labourer who transfers to the building industry.

The government's first task must be to improve occupational mobility. Entry into certain occupations should be made less difficult, e.g. by giving information on opportunities in other industries and occupations and by persuading trade unions to relax their apprenticeship rules. More important, people must be trained in the new skills required by expanding industries, e.g. through local Training and Enterprise Councils.

Improving the geographical mobility of workers to the more prosperous regions operates chiefly under the government's Employment Transfer Scheme. This consists of granting financial aid towards moving costs, providing informa-tion on prospects in other parts of the country and giving free fares to a place of work away from the home town.

21.3 Work to the workers

While a 'workers-to-the-work' policy has a role to play in correcting regional imbalance, it suffers from: (i) an exclusive concern with unemployment to the

neglect of other consequences of regional imbalance; (ii) a failure to recognise the macro effects of the outward movement of workers.

Thus taking work to the workers is now regarded as the policy most likely to effect a long-term solution to the problem for it reduces regional differences in income and the rate of growth as well as in unemployment. By helping the more immobile workers, such as older people and married women, it stimulates the activity rate. It also avoids forcing workers to leave areas to which they are attached, relieves the growing congestion in south-east England, and prevents the loss of social capital resulting from the depopulation of depressed areas. Above all, it works in harmony with Keynesian macro theory. The 'multiplier' operates for regional economies in much the same way as it does for the national economy. Moving unemployed workers and their families reduces spending in the area (e.g. because unemployment benefits are no longer being drawn) and this gives rise to a negative multiplier. In contrast, moving firms into the area generates spending power and produces a positive multiplier, variously calculated at between 1.25 and 1.50.

On the other hand, a policy of locating firms in depressed areas may involve them in higher costs. Their desire to establish plants in the south-east is to secure location advantages, such as supply of skilled workers, easier and less costly communications, contact with complementary firms and nearness to EC markets. The government, therefore, has to offer financial grants to induce firms to establish or expand in Assisted Areas (Figure 21.1). These were revised in August, 1993 to take account of new areas and localities where unemployment problems were significant because of major closures (e.g. Mansfield – coal-mines) or even in the south of England, through the contraction of defence or holiday demand (e.g. Portland, Thanet, Isle of Wight). Assisted Areas now cover some 34 per cent of the working population of Great Britain.

In these Areas *Regional Selective Assistance*, amounting to £300 mn in 1993, is given on a discretionary basis, mainly through *Project Grants* based on the capital cost and the number of jobs created.

Regional Enterprise Grants are also available in the Assisted Areas designated as Development Areas and in other areas covered by EC schemes. Such grants are to support investment and innovations by firms having fewer than 25 employees.

In addition to these grants, other incentives include:

  (i) help for transferring key workers;
 (ii) factories for renting on favourable terms, such as an initial rent-free period;
(iii) loans at favourable rates of interest;
(iv) grants towards the cost of reclaiming derelict land in certain designated areas.

All the above are in addition to the investment incentives (through tax concessions) available to industry throughout the whole country.

In the dispersal of industry the government has set an example. Thus the Department of Health and Social Security is centred in Newcastle, the Girobank in Bootle, and the Driving and Vehicle Licensing Centre in Swansea.

**Figure 21.1   Assisted Areas, GB: as defined by the Department of Trade and
          Industry at 1 August 1993**

21.4   Regional planning

While development area policy deals with special areas needing extra help, there is
now a trend towards planning for larger regions each with an infrastructure which is

attractive to industry and management in that it features: good executive housing; educational, cultural and recreational facilities; first-class shops and hotels; and a modern road network, rail services and airport. The aim is to secure a broad-based industrial structure, with special consideration being given to service industries.

The whole country is divided into ten regions (eight for England and one each for Scotland and Wales). Each region has an Economic Planning Board, consisting of civil servants from the main government departments concerned with regional planning, whose task is to formulate plans and coordinate the work of the various departments.

Primary responsibility for regional development lies with the Department of Trade and Industry which has an Industrial Development Unit to help with the appraisal, negotiation and monitoring of projects.

21.5   Inner city regeneration

The concentration of policy on inter-regional differences has tended to divert attention away from the problems existing *within* regions. In many inner city areas, the failure to replace the older industries, as they have declined, with commerce or new industry has left a picture of physical dereliction and decay. Moreover, because their workers have not the necessary skills or cannot afford to emigrate to the greenfield sites of the new industries, inner city areas often suffer a high rate of unemployment. Thus Greater London's rate of 11.6 per cent (April 1993) is only exceeded by those of the North and N. Ireland regions, while certain inner districts e.g. Poplar, have a rate of up to 16 per cent.

With the object of regenerating these run-down areas, the government has introduced a variety of schemes. Beginning in 1981, it has established *Enterprise Zones* (currently 18) in such places as London docklands, Swansea, Manchester, Clydebank, Belfast and Hartlepool. Each zone covers up to 5000 acres and for ten years firms there enjoy special benefits, e.g. exemption from business rates, simplified planning procedures and favourable taxation on capital spending. This was followed by *Derelict Land Grants* to enable local authorities to bring derelict land back into use for industrial and commercial development.

Urban Development Grants and Urban Regeneration Grants (now combined in a *City Grant*) were introduced in 1987. Grant applications are made to the Department of the Environment by local authorities on behalf of developers. The object is to encourage new private sector capital investment in urban areas which have suffered physical degeneration and unemployment through the decline of traditional industries.

Regeneration in many major cities is the task of Urban Development Corporations, e.g. the London Dockland Development Corporation. They have been given powers to buy and manage land, by-pass formal planning procedures, and provide buildings and infrastructure in order to attract private sector investment in their areas.

A recent innovation has been the '*City Challenge*'. Here the Department of the Environment awards grants to local authorities who compete for a share of the global sum available on the merit of the scheme submitted.

At present (1993) an *Urban Regeneration Agency* is being set up to combine three functions: City Grant and Derelict Land Grant from the Department of the Environment and English Estates from the Department of Trade and Industry. It will be able to assemble and reclaim land, authorise grants and make compulsory purchase orders.

21.6   Regional policy in the context of the EC

A healthy integrated Community – at both economic and political levels – is possible only if progress is made towards reducing disparities in economic opportunity between regions within the Community. Indeed, while the foregoing reasons for regional policy are all relevant at the Community level, additional considerations apply:

(a) Physical controls are more difficult to apply in the EC context. Not only are they at variance with the objective of greater mobility within the EC, but firms have the option of relocating in a prosperous region of another member state.

(b) The depressed peripheral regions of Scotland, Northern Ireland, southern Italy, etc., are more distant from the expanding centre of the Community – south-east England through to north-east France and western Germany – than they are from the centres of their own countries. This Community 'centre' forms a concentrated market to which industries are likely to be increasingly attracted, thereby adding to its dominance.

(c) The EC embraces regions exhibiting wider economic disparities than in any one member-state. Moreover, regional problems are more heterogeneous – for example, whereas the UK depressed regions are mainly industrial, Italy has many depressed agricultural areas.

These additional considerations mean that the formulation of an effective EC regional policy is a difficult task. Not only must it respond quickly as new regional problems arise, but it has to be linked with, and be complementary to, the individual nation's regional policy. Indeed EC policy should also coordinate the regional policies of member-states; for example, a physical control in one country must not be undermined by a firm being able to locate in another country.

It follows, therefore, that regional policy must be handled to a substantial degree at the Community level and be wide-ranging in the measures employed so that one reinforces the others. Above all, to achieve greater equity, it must envisage substantial transfers of income through incentive funds which are in addition to those provided by the member states.

The emphasis of the EC's Structural Fund is now on regional development *programmes* rather than on individual projects. It can insist that grants are actually spent in the specified region.

The UK is a major beneficiary of the Fund receiving 38.3 per cent of EC regional aid 1989–93.

The unemployed of the depressed areas also benefit from aid from the Social Fund, while loans are available on favourable terms from the European Investment Bank.

22 Growth

22.1 The nature of growth

(a) The meaning of 'growth'
When there are unemployed resources, the economy's *actual* output is below its *potential* output; in terms of Figure 22.1 the economy is producing inside the production possibility curve, say at point *A*. Here output can be increased, even in the relatively short term, by measures which absorb unemployed resources.

But, by itself, full employment of an economy's resources does not necessarily mean that the economy will grow. Growth is essentially a 'long-run phenomenon – the *potential* full-employment output of the economy is increasing over time. Whereas full employment simply means that the economy is producing on a point on the production possibility curve I, growth means that, over time, the curve is pushed outwards to II and III. Even with full employment of resources, advanced economies can achieve an *annual* growth rate of 3 per cent.

Increases in the productive capacity in the economy over time are usually measured by calculating the rate of change of real gross national product per head of the population (see pp. 220). However, when people talk about 'growth' they are thinking chiefly of the difference it makes to the standard of living rather than to output itself. Allowances have to be made, therefore, for the defects of GNP as an indication of the standard of living (see pp. 218). Also, when measuring long-term *secu-*
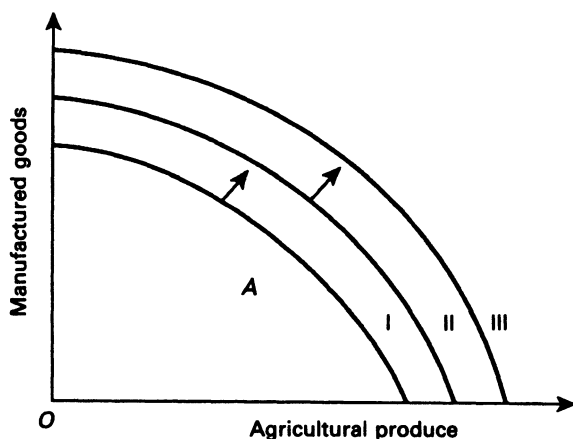
Figure 22.1 Economic Growth

lar growth by the change in real GNP per head between different years, we must recognise that unemployment rates from which measurement starts may differ.

(b) Advantages of growth

Economic growth is the major factor for achieving improvements in the standard of living – more consumer goods, a shorter working week, and so on. While such improvements occur almost imperceptibly from year to year, small differences in the *annual rate* of growth produce large differences in the *speed* of growth. For instance, a rate of growth of 2.5 per cent per annum will double real GNP in twenty-eight years, whereas a 3 per cent rate doubles GNP in only twenty-four years.

In addition, growth makes it easier for the government to achieve its economic policy objectives. Revenue from taxation increases, allowing government services, e.g. health care, to be expanded without raising the *rates* of tax. Income can also be redistributed in favour of the poorer members of society while still allowing the standard of living of the better-off to show some improvement.

However, economic growth does have its costs (see pp. 263)

22.2 Achieving growth

(a) Factors producing growth

There are five basic causes of growth:

(i) *A rise in the productivity of existing factors*

In the short run, productivity may be raised by improvements in organisation, which secure, for example, more division of labour and economies of large-scale production, or a more intensive use of capital equipment (e.g. the adoption of shift-working). Physical improvements for the labour force, e.g. better food and working conditions, may also increase productivity.

In the longer run, more significant increases can come with education and the acquisition of skills through training. These really represent, however, an increase in the capital invested in labour.

It is also important to draw attention to the differences in personal incentives provided by the market economy and the command economy. Compare, for instance, the growth rates of Hong Kong and Poland over the last forty years.

(ii) *An increase in the available stock of factors of production*

(1) A RISE IN THE LABOUR INPUT

The size of the labour input can increase relative to the total population through either an increase in the number of hours worked per worker, or an increase in the ratio of the working population to the total population. The first is hardly likely to be a cause of growth in normal conditions, for as living standards improve the tendency is to demand more leisure. The second, however, may come about if the percentage of the population in the working age group increases or if attitudes to work change (see p. 134).
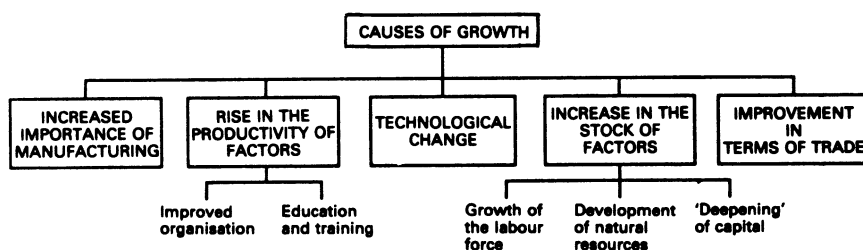
Figure 22.2 Factors leading to growth

(2) DEVELOPMENT OF NATURAL RESOURCES

North Sea natural gas and oil, for instance, have allowed Britain to obtain her fuel supplies from fewer factors of production, enabling resources to be transferred to other output and thus promoting growth.

(3) ADDITIONAL CAPITAL EQUIPMENT

Here we must distinguish between 'widening' and 'deepening' capital. Widening capital – adding similar capital equipment – is necessary if the labor force increases, in order to maintain the existing capital–labour ratio and thus output per head. Suppose 10 men, digging a long ditch, have 5 spades between them. If the labour force is increased to 20 men the capital–labour ratio falls from 1:2 to 1:4 unless 'widening' takes place – that is, unless another 5 spades are provided to maintain the existing ratio. 'Widening' does not increase productivity; it simply prevents diminishing returns to labour setting in.

'Deepening' capital occurs when the capital–labour ratio is increased. If, for example, when there were 10 spades to 20 men, the men were given a further 10 spades, the capital–labour ratio would be raised to 1:1.

(iii) *Technological change*

All we have done in our example so far has been to increase the stock of a given kind of capital equipment, spades. Over time, however, productivity can be raised more significantly by technological improvements. Thus the twenty men and their spades may be replaced by a single trench-digger and its driver. Because this does the job more quickly and efficiently the remaining nineteen men are released for other kinds of work.

In practice, all three causes are usually operating at the same time to increase productivity. Thus, as the labour force or natural resources are expanded, new capital is required, and this allows for the introduction of new techniques.

The speed with which new capital and improvements are introduced also depends upon the price of capital equipment relative to the wages of the labour for which it can be substituted. Over the last fifty years, wages have tended to rise relative to the cost of capital equipment. The effect has been to increase the rate of technological change in such industries as agriculture, cargo handling, transport, shipbuilding and mining.

Other factors affecting growth are:

(iv) *Fundamental changes in the composition of the national output*
As a country's standard of living improves so spending switches from agricultural to manufactured goods and then towards services. Since the opportunities for employing more capital and technical improvement are greatest in manufacturing, the growth rate increases as countries industrialise but then slows down as the relative demand for personal and government services increases.

(v) *A sustained improvement in the terms of trade* (see p. 208)

(c) Constraints upon growth
In practice the UK has not succeeded in sustaining a 3 per cent annual growth rate. Thus from 1960 to 1993 it averaged only 2.3 per cent. Why is this?

First, when employment was buoyant, consumer demand left fewer resources available for investment.

Second, inflation has proved inimical to investment (see p. 240). Government disinflationary measures undermine the confidence of entrepreneurs.

Third, at times growth has been incompatible with other government objectives. For instance, 'stop' policies have been necessary because, as the economy expands, increasing imports produce balance-of-payments difficulties (see p. 241).

Fourth, growth entails costs additional to the reduced current consumption necessary to accumulate capital. Growth usually requires change, and the more rapid the growth, the greater the change. Changes in the structure of the economy are, as we have seen, bound to lead to some unemployment, and if growth is to be achieved people must be willing to change jobs quite radically, three or four times in their working lives. This will entail retraining and probably moving around the country.

Fifth, growth is not achieved without environmental costs – pollution, noise, loss of natural beauty, destruction of wildlife habitat. And, as material wealth grows, people are inclined to question its full costs, preferring some slowing down in the growth rate.

Finally, and on the same theme, Britain's growth has been taken in forms not appearing in GNP calculations (see p. 219). Increasing welfare may mean preferring a quiet life and more leisure to the 'rat-race' and stress of accumulating material goods.

22.3   The government and growth

(a) Dificulties in framing policy
In framing policy for a satisfactory rate of growth, the government faces difficulties.

(1) How is growth to be measured? We have already drawn attention to the ambiguities of GNP figures as a measure of improved living standards. Is growth to take the form of more leisure, less spending on defence, a litter-free environment, the preservation of town and rural beauty, safer and more comfortable travel?

(2) Of all the factors producing growth, which plays the greatest part – education of the population, training of labour, additional capital equipment, technological advance? There is no real means of measuring.

(3) Growth does not proceed at a *steady* rate but rather by unavoidable take-offs and slow-downs around an upward trend, and firms are generally able to cope, e.g. by adjusting their stock levels. However, if these fluctuations are too wide, uncertainty can deter investment.

(b) Government policy

Since there is no single satisfactory theory of growth, government action seeks to promote the factors essential to growth.

Growth results mainly from capital investment. But accumulating capital involves foregoing present consumption – saving. Thus a first requirement for the government is to ensure that its policies, especially as regards taxation, will provide the real saving required.

Second, additions to capital take place in both the public and private sectors. The government itself is largely responsible for investment in the public sector – the infrastructure (such as roads, hospitals, schools), education, and the training of the labour force.

On the other hand, the scale of the government's priorities must not 'crowd out' desirable private sector investment. Decisions here are based on expected profitability. Private investment can be increased, therefore, by providing a stable economic background free from 'stop-go' policies so that fixed capital formation, research and development (R & D) are not inhibited.

While we must not ignore micro supply-side measures, the major spurts in growth have come through breakthroughs in technology by innovation and the application of inventions. Most of these, e.g. aircraft, computers, antibiotics and other drugs, plant-breeding, animal selection, pesticides, etc. are the result of long-term R & D. While about a half of R & D is carried out in the private sector, the other half, particularly as regards defence, is largely through government-sponsored research bodies. Britain's civil R & D has tended to lag behind other countries, but with the ending of the Cold War some transfer from defence is likely.

Finally, the government has to encourage the application of the fruits of R & D so that innovation and inventions are transformed into new marketable products.

⟨23⟩ Public finance

23.1 The distribution of income

Public finance is concerned with government spending and revenue, the difference between them (the PSBR), and their magnitudes relative to GDP.

The government is interested in the distribution of income and wealth between its people for reasons of fairness, social harmony and its effect on the macro variables, e.g. saving. And, although the welfare effects of redistribution cannot be measured objectively, the nature of government expenditure and how revenue is raised do effect a redistribution of income.

Thus a sample of households taken by the Family Expenditure Survey in 1989 showed that while on average the bottom one-fifth started off with an original yearly income of £1280, cash benefits less direct taxes and rates gave them a disposable income of £3960. Indirect taxes took £1120 of this, but benefits in kind (e.g. education, national health services, housing and travel subsidies) produced a final income of £5200.

In contrast the top one-fifth had an original income of £30,050 and, after making the same adjustments, had a disposable income of £23,970 and a final income of £21,400.

Thus, although considerable inequality of incomes still persists, government policies do serve to reduce the disparity. And even though decisions on redistribution are politically subjective, the economist can still draw attention to both the likely immediate and long-term economic results of redistribution policies and thus lay bare the implications of the different choices open to the government.

23.2 Government expenditure

(a) Limits to government spending
Today government spending takes about 45% of the gross national product – a remarkable increase since 1910, when the figure was only about 10%. The government is now spending on a much wider range of activities.

Many items of government spending, e.g. pensions, national-debt interest and grants to local authorities, are unavoidable since by nature they are basically contractual. It may therefore seem that the government has merely to estimate its expenditure and impose taxes to cover it. But this is not the case.

Since goods and services in the economy as a whole are limited, the government has to cut its coat according to its cloth, asking such questions as: What can be afforded for the Arts Council? How much can be given to local authorities? Can university education be expanded? Can national-insurance contributions be

reduced? The economic problem, involving decisions at the margin, confronts private persons and the government alike. When there is full employment, the government can only secure more goods and services by allowing the private sector less. In the last resort the division rests on a political decision.

(b) The distribution of government expenditure
Government spending can be classified under the following headings:
  (i)   *Defence*, which has accounted for about one-eleventh of all government spending.
 (ii)   *Internal security* – the police, law-enforcement and fire brigades.
(iii)   *Social responsibilities* – education, and protection against the hazards of sickness, unemployment and old age.
(iv)   *Economic policy,* covering subsidies to agriculture and industry, help for Assisted Areas, worker training and the provision of capital to the nationalised industries.
 (v)   *Miscellaneous,* including expenditure on diplomatic services, grants to local authorities and – the largest single item – interest on the national debt.

(c) How government expenditure is financed
In the same way that firms have to pay for both variable and fixed factors, the government has to spend not only on single-use goods and services but also on goods which render services over long periods. The first, which involve regular yearly spending, should be paid for out of regular yearly income. But capital spending, on such items as roads, loans to nationalised industries and university building, is more fairly financed by borrowing, for the repayment of the capital then partly falls on future beneficiaries.

Regular yearly income comes from two main sources: (i) miscellaneous receipts, chiefly interest on loans, rents and charges on goods and services (such as medical prescriptions): and (ii) taxation, described in more detail below.

The difference between expenditure and revenue has to be covered by borrowing – the PSBR. Yearly borrowing increases the National Debt, the capital sum of accumulated borrowing.

Government borrowing takes the form of:

(i) Short-term loans from the sale of Treasury bills. Originally these were used to bridge the time-gap between expenditure and receipts from taxation, but – because it is cheaper to borrow short than long – they eventually became a major means of government borrowing. Nevertheless, being liquid assets to the banks, there are possible inflationary effects.

(ii) Medium- and long-term loans are obtained by selling stock having a minimum currency of five years. Some, such as $3\frac{1}{2}$ per cent War Loan, are undated.

(iii) '*Non-market' borrowing,* through National Savings Certificates, Premium Savings Bonds, etc. and the National Savings Bank.

Since 1980 additional capital funds have been obtained by the 'privatisation' sales of publicly owned assets (see p. 172).

REVENUE
Where each £ comes from

EXPENDITURE
How each £ was spent

Revenue		Expenditure	
Income taxes	26	Social security	30
National insurance and other contributions	17	Defence	9
Corporation tax	7	Health and personal social services	15
Capital taxes	2	Education	12
Council tax and uniform business rate	9	Housing and environmental services	6
Value added tax	15	Transport	4
Road fuel, alcohol and tobacco duties	10	Law, order and protective services	5
		Trade, industry, energy and employment	3
Other sources	5	Other	9
Borrowing	9	Debt interest	7

Totals of revenue and expenditure £269 billion

Figure 23.1   Public income and expenditure, 1992–3

23.3 The modern approach to taxation

(a) Taxation and government policy

Until the end of the nineteenth century, the functions of the state were mainly defence and the preservation of law and order, and taxation was levied primarily for revenue purposes.

To meet the vast increase in government spending over the last fifty years, higher rates of taxation have been imposed and new taxes introduced. As we shall see, these additions have provided new means for promoting economic and social policies. Briefly, by its fiscal measures, the government can: (i) influence the allocation of resources, e.g. to allow for social costs and benefits; (ii) exercise an overall control of the economy, mainly to promote employment or check inflation; (iii) promote economic growth, e.g. through tax allowances for investment and subsidies to agriculture; (iv) re-distribute income and wealth, e.g. by providing welfare benefits and community, public and subsidised merit goods, and by progressive taxation.

(b) The attributes of a good tax system

In his *Wealth of Nations 1776*, Adam Smith was able to confine his principles of taxation to four simple tenets: persons should pay according to their ability; the tax should be certain and clear to everybody concerned; the convenience of the contributor should be studied as regards payment; and the cost of collection should be small relative to yield.

While today the main purpose of any tax is usually to raise money, emphasis is now placed on other attributes. Indeed, the ideal tax should be:

(i) *Productive of revenue*

All taxes cost money to collect and are unpopular. The yield of any tax should therefore at least cover the cost of collection, with something to spare to offset the vexation caused. In practice, too, a single tax with a high yield is better than a number of taxes each having a small yield, for these complicate the tax structure and make it difficult to understand and administer. The Chancellor should also be able to estimate the yield from a tax if the budget is to be used for adjusting overall demand.

(ii) *Certain to the taxpayer*

Not only should a taxpayer know exactly when and where he has to pay his tax, but he should find it difficult to evade payment. Indirect taxes score heavily here.

(iii) *Convenient to the contributor*

Bad debts and evasion are reduced if the time and manner of tax payment are related to how people receive and spend their incomes.

(iv) *Impartial between one person and another*

All persons similarly placed should pay the same tax. Thus, while non-smokers do not pay the selective tax on tobacco, all smokers do.

   Yet, although there is impartiality in this sense, the concentration of indirect taxes on a few goods – chiefly tobacco, alcoholic drink and motoring – severely penalises certain forms of spending. One of the objects of introducing VAT (see p. 272) was to broaden the tax base.

(v) *Adjustable*
A tax should be capable of variation, both up and down, according to changes in policy.

(vi) *Automatic in stabilising the economy*
Varying the relationship between government expenditure and revenue is one of the major devices for stabilising the economy especially as regards employment and the price level. Through tax changes, the Chancellor of the Exchequer can vary the purchasing power of the community. Usually he has to make a deliberate adjustment in his budget, but it is helpful if taxes operate automatically in the desired direction. Thus, when money income increases, so do income tax and VAT yields, thereby reducing inflationary pressure, and vice versa.

(vii) *Harmless to effort and initiative*
This becomes important as direct taxation increases. High rates of income tax, for instance, may induce the taxpayer to take his income in the form of leisure or reduce his willingness to undergo training or seek promotion. The extent to which this occurs, however, is uncertain. If a person has fixed money commitments, e.g. hire-purchase instalments, mortgage repayments and insurance premiums, he may have to work harder to meet them when his income is reduced. Furthermore, if we assume that a high rate of income tax is a disincentive to effort, we infer that people always look upon work as distasteful, while leisure is seen as a pleasurable alternative. This may be true for many, but in the high-income brackets there are some who find their work enjoyable. Last, most workers have to work a 'normal' week, and can only vary their hours as regards overtime.
   The disincentive effect is more likely to occur when there is a sudden jump in the rate of tax between one income level and another. People reduce their effort at the higher marginal level.
   High direct taxes can also affect enterprise and efficiency. By eroding the wage differentials between skilled and unskilled labour they reduce incentives. Similarly, entrepreneurs are only prepared to accept risks if the rewards are commensurate.
   Furthermore, high taxation of profits and income means that the penalty of inefficiency is not borne entirely by the taxpayer. Because income is smaller, less tax is paid and so a part of the cost falls on the government.

(viii) *Consistent with government policy*
Although the tax structure should not change frequently, individual taxes must be constantly reviewed to promote current government policy. To encourage effort, should income from work be taxed at a lower rate than investment income? Will an indirect tax, by raising the cost of living, increase wage-push inflation?

Indirect taxes can be adapted to specific objectives, e.g. cigarettes bear a high selective tax, while exports are 'zero-rated' for VAT (see below).

(ix) *Minimal in its effect on the optimum allocation of resources*

An indirect tax on a *particular* good results in resources' not being fully allocated according to the preferences of consumers. This may be acceptable where social costs and benefits have to be allowed for. But normally, to maximise satisfaction, consumers spend so that the marginal utility relative to price is equal for all goods. This relationship is destroyed by a tax on one good for this raises its price, resulting in a redistribution of consumers' expenditure and thus of the factors of production. In addition, imposing a tax on the product of an industry may have long-term harmful effects (see p. 276).

Again, selective indirect taxes entail a greater loss to the consumer than an income tax yielding the same amount. Unlike the latter, selective indirect taxes change the relative prices of goods, so that consumers have to re-arrange their patterns of expenditure. This substitution involves a loss of satisfaction in addition to that suffered through the reduction in income.

Finally, direct taxes can affect the supply of factors, particularly capital, to industry. High taxation may discourage saving; it certainly reduces the power to save. This is not serious for large companies, who can borrow on the open market. But the major sources of capital for small businesses are the owners' personal savings and profits ploughed back. Thus income tax and corporation tax deprive small and risky but often progressive companies of much-needed capital.

Not only that, but high direct taxes may repel foreign capital. Although the deduction of income tax on dividends may be refunded, the company still has to bear corporation tax on profits (at 33% in 1993). The amount available to shareholders is therefore less, and the declared dividend correspondingly smaller. Consequently people may prefer to invest in companies operating in countries where there is a higher return to capital – a higher return which is the result not of superior efficiency but simply of lower taxes.

(x) *Equitable its distribution of the tax burden*

Taxes can be classified according to the proportion of a person's income which is deducted:

(1)   A *regressive* tax takes a higher proportion of the poorer person's income than of the richer. Indirect taxes which levy a fixed sum irrespective of income, e.g. television licences, are regressive.

(2)   A *proportional* tax takes a fixed proportion of one's income. Thus income tax is proportional between £3000 and £23,700 of taxable income, 25% of every £1 being taken in tax.

(3)   A *progressive* tax takes a higher proportion of income as income increases (Figure 23.2). Income tax is progressive in that the first £3000 of taxable income is taxed at 20%, then up to £23,600 at 25%, and above this at 40%.

Justification for taxing the rich man more highly than the poor man rests on the assumption that the law of diminishing utility applies to additional income, and
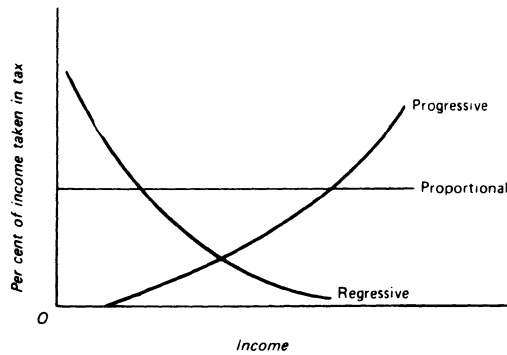
Figure 23.2 The difference between regressive, proportional and progressive taxes

that an extra £50 affords less pleasure to the rich man than to the poor man. Thus taking from the rich involves less hardship than taking from the poor. Generally this can be accepted as true, but we can never be sure, simply because there is no absolute measure of personal satisfaction.

23.4 The structure of taxation

Because the objectives of taxation are now so varied and may even be incompatible, no single tax is completely perfect. Consequently there must be a structure of taxation, combining different taxes which can be varied according to changes in emphasis on different objectives.

The following classification of taxes is based on their methods of payment.

(a) Direct taxes
With these taxes the person makes payment direct to the revenue authorities – the Board of Inland Revenue or the local authority. Usually each individual's tax liability is assessed separately.

(i) *Income tax*
'Taxable income', which is subject to a basic rate (now 25%) between £3000 and £23,700 is arrived at after allowing deductions depending on marital status and other personal circumstances.

(ii) *Corporation tax*
All profits, whether distributed or not, are taxed at the same rate (33% in 1994). A part, 20 per cent, is imputed to shareholders and paid in advance when the dividend is paid. This advance payment is allowed against the mainstream 33% corporation-tax payment (which is always paid in arrears), while for the shareholder it counts as a 'tax credit', and is refundable if income tax is not paid because of low income.

(iii) *Capital gains tax*
A tax is now levied at the individual's income tax rate of any capital gain when an asset is disposed of. Owner-occupied houses, cars, National Savings Certificates and goods and chattels worth less than £6,000 are excluded, and losses may be offset against gains.

Where the gain does not exceed £5,800 in any year, no tax is payable.

(iv) *Inheritance tax*
Inheritance tax applies to lifetime gifts as well as to legacies, though the former generally bear only half the latter's rate of tax. The starting point is £150,000 and the rate of duty is 40%. Gifts made more than seven years before death are exempt from tax.

(v) *Other taxes*
These consist of stamp duties (payable on financial contracts), motor-vehicle duties and a Petroleum Revenue Tax.

The Council Tax and Uniform Business Rate levied on business premises, can also be regarded as direct taxes.

Direct taxes yield nearly two-thirds of total revenue. Their great merit is that, being progressive and assessed according to the individual's circumstances, they ensure that the heaviest burdens are placed on the broadest backs. Their progressive character also gives additional weight to their role as built-in stabilisers.

Their main disadvantage is that when the rate of tax is high there may be disincentive effects. As a result indirect taxes also have to be levied.

(b) Indirect taxes
Indirect taxes on goods and services are so called because the revenue authority (the Department of Customs and Excise) collects them from the seller, who, as far as possible, passes the burden on to the consumer by including the duty in the final selling price of the good (see p. 276). They may be *specific* (i.e. consisting of a fixed sum irrespective of the value of the good) or *ad valorem* (i.e. consisting of a given percentage of the value of the good).

Indirect taxes may be divided into:
  (i) Customs duties levied at EC rates on goods imported from countries outside the EC.
 (ii) Excise duties on home-produced goods and services, e.g. beer, whisky, petrol, cigarettes and gambling.
(iii) Value Added Tax (VAT): an *ad valorem* tax, levied on most goods and services at each stage of production at a basic rate. Using Figure 8 as an example, a VAT at $17\frac{1}{2}$ per cent paid by the consumer on the table in the shop would be £17.50 making a total purchase price of £117.50. The VAT, however, would have been paid at each stage of production: tree-grower £5.25; saw-miller £3.50; table manufacturer £5.25; retailer £3.50. In practice each producer would pay to the Customs and Excise the full $17\frac{1}{2}$ per cent tax of the goods as invoiced by him less the VAT paid by his suppliers of materials, etc. as shown on their invoices. Thus the

retailer would pay the Customs and Excise £3.50, i.e. £17.50 minus the
VAT of £14 charged to him.

Some goods, e.g. food (except meals out), houses, books, newspapers,
public-transport fares, medicines on prescription, etc. are zero-rated. This
means that the final seller charges no VAT *and* can reclaim any VAT
invoiced by intermediary producers. Other goods, e.g. rents and medical
services, are 'exempt'. Here no VAT is charged by the final seller, but any
VAT paid by an intermediary, e.g. a landlord paying for building repairs,
cannot be reclaimed.

The main merit of VAT is that it is broad-based, the yield increasing
almost proportionately to consumer spending. Moreover, since VAT
covers most forms of spending, it does not distort consumer choice as
much as a highly selective tax.

On the other hand it can be argued that a general tax on spending is regressive,
for it hits harder those on lower incomes, since they spend proportionately more
of their incomes. This fact is tempered somewhat, however, by zero-rating goods
regarded as necessities.

Indirect taxes give a certain and often an immediate yield and can be adjusted
to specific objectives of government policy. On the other hand, by being
regressive they undo some of the re-distributive effects of direct taxes.

(c) The distribution of the burden of an indirect tax between consumers and producers

When a good is subject to a selective tax this does not mean that its price will rise
by the full amount of the tax. Consider the following demand and supply
schedules for commodity X.

Price of X (pence)	Demand ('000 lb.)	Supply ('000 lb.)
12	60	150
11	70	130
10	80	110
9	90	90
8	100	70

The equilibrium price is 9p. Now suppose a tax of 3p is charged on the producer
for each unit of X he puts on the market. This means that whereas before the tax
he supplied 70,000 units at a price of 8p, he will now only supply this quantity at
11p (because 3p would go in tax). Similarly, 90,000 units will only be supplied at
a price of 12p instead of 9p. Thus the effect of the 3p tax can be shown by the
shift in the supply curve from S to S_1 (Figure 23.3). This gives a new equilibrium
price of 11p, the buyer paying 2p more and the producer receiving 1p less, the
quantity traded falling from 90,000 units to 70,000 units.

The result is the same if the 3p tax is levied on purchasers. Before the tax,
10,000 units were demanded at a price of 8p. If purchasers now have to pay a 3p
tax, this is equivalent to a price of 11p including tax, and so they will demand
only 70,000 units. Similarly, for a price of 9p they will demand only 60,000 units
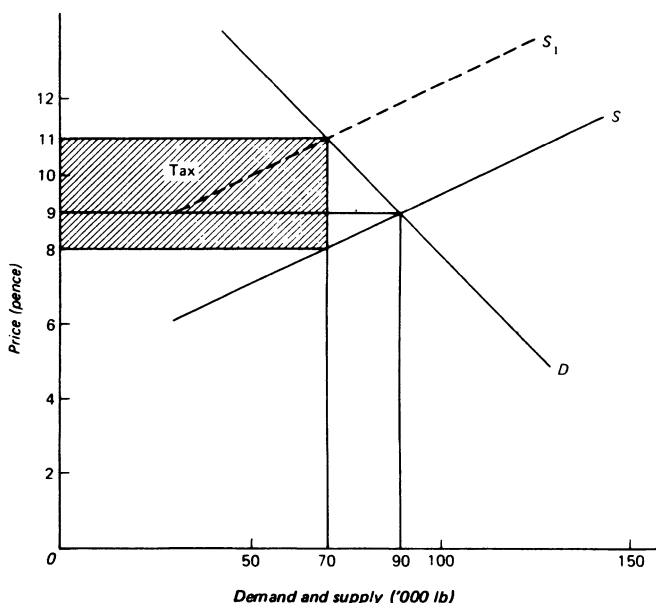
Figure 23.3 The diagrammatic representation of a tax on the supply side

instead of 90,000. Thus the effect of the 3p tax imposed on buyers can be shown by the move in the demand curve from D to D_1 (Figure 23.4). This gives a new equilibrium price of 11p (8p at which the market is cleared, plus 3p tax), and the quantity traded falls from 90,000 to 70,000.

The amount of the tax falling on consumers as compared with that falling on producers is directly proportional to the elasticity of supply as compared with the elasticity of demand. That is:

$$\frac{\text{consumers'share of tax}}{\text{producers' share of tax}} = \frac{\text{elasticity of supply}}{\text{elasticity of demand}}$$

That this proposition is likely to be true can be seen from the following argument. When a tax is imposed, the reaction of the producer is to try to push the burden of the tax on to the consumer, while similarly the consumer tries to push it on to the producer. Who wins? Simply the one whose bargaining position is stronger. This will depend upon the ability to switch to producing substitutes if the price falls as compared with the ability to switch to buying substitutes if the price rises. Now the possibility of substitution largely determines elasticities of supply and demand. Thus the relative burden on the tax paid by producers and consumers depends upon relative elasticities of supply and demand.

(d) The effect of an indirect tax on the size of an industry

The greater the elasticities of demand and supply, the greater will be the effect of a tax in reducing output. We can show this diagrammatically (Figures 23.5 and 23.6).
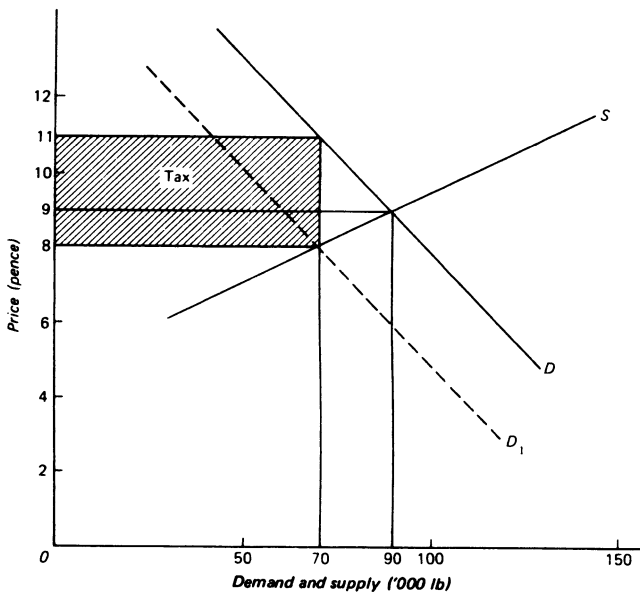
Figure 23.4 The diagrammatic representation of a tax on the demand side

(i) *Elasticity of demand*

Before the tax is imposed, total output is *OM* (Figure 23.5). The effect of the tax is to raise the supply curve from *S* to *S₁*. Two demand curves are shown, *D*ₐ being less elastic than *D*ᵦ at price *OP*. The effect of the tax is to reduce output to *OM₁* where demand is *D*ₐ, and to *OM₂* where it is *D*ᵦ. In the latter case consumers are more able to switch to buying substitutes.
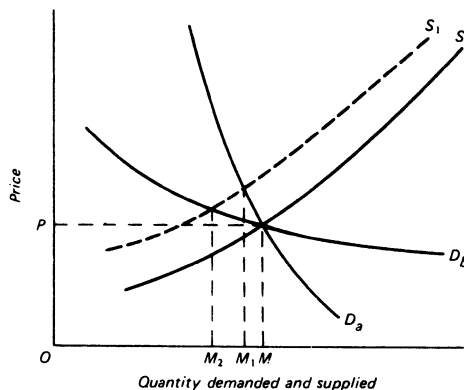
Figure 23.5 The relationship of elasticity of demand and production when a tax is imposed on a good

(ii) *Elasticity of supply*

Before the tax is imposed, total output is *OM* (Figure 23.6). The effect of the tax is to lower the demand curve from *D* to D_1. Two supply curves are shown, S_a being less elastic than S_b at price *OP*. The effect of tax is to reduce output to OM_1 where supply is S_a, and to OM_2 where it is S_b. In the latter case producers are more able to produce alternative goods.
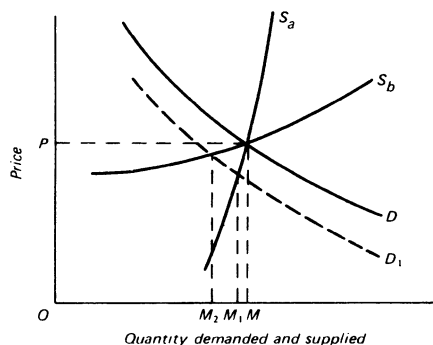
Figure 23.5   The relationship of elasticity of supply and production when a tax is imposed on a good

   This proposition has important practical applications. (i) The government may use a subsidy (which can be illustrated by moving the supply curve to the right) to increase the production, and thus employment, of an industry. The effect will be more pronounced where demand and supply are elastic. (ii) Because the effect of a tax is to reduce production, even a temporary tax may harm an industry. This is particularly so where home demand is elastic and production takes place under decreasing costs, for the smaller demand will raise export prices. Thus a selective tax on cars would not only reduce home demand but also, by doing so, lose economies of scale, thereby putting up prices in both home and foreign markets. Even when the tax is subsequently withdrawn, foreign markets may not be regained.
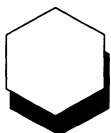
Part VIII
International trade

⟨24⟩ **International trade**

24.1 Why international trade?

(a) How international trade arises

International trade arises simply because countries differ in their demand for goods and in their ability to produce them.

On the demand side, a country may be able to produce a particular good but not in the quantity it requires. The USA, for instance, is a net importer of oil. On the other hand, Kuwait does not require all the oil she can produce. Without international trade most of her deposits would remain untapped.

On the supply side, resources are not evenly distributed throughout the world. One country may have an abundance of land; another may have a skilled labour force. Capital, oil, mineral deposits, cheap unskilled labour and a tropical climate are other factors possessed by different countries in varying amounts.

Nor can these factors be transferred easily from one country to another. Climate, land and mineral deposits are obviously specific. Labour is far more immobile internationally than within its own national boundaries. Capital, too, moves less easily; exchange controls, political risks and simple ignorance of possibilities may prevent investors from moving funds abroad.

Because factors are difficult to shift, the alternative – moving goods made by those factors – is adopted. What happens is that countries specialise in producing those goods in which they have the greatest comparative advantage, exchanging them for the goods of other countries.

(b) The advantages of international trade

(i) *It enables countries to obtain the benefits of specialisation*
Specialisation by countries improves their standard of living.

(1) It is obvious that without international trade many countries would have to go without certain products. Britain, for instance, has no gold or aluminum, and Sweden no oil.

(2) More important, many goods can be enjoyed which if produced at home would be available only to the very wealthy, for instance bananas, spices, oranges and peaches in Britain. But this benefit can be applied generally to all imports. The 'law of comparative costs' shows that, provided countries differ in the relative costs of producing certain goods, they can probably gain by specialisation and trade.

Suppose that there are two countries, A and B, each producing just two commodities, wheat and cars. Each has the same amount of capital and the same number of labourers, but A has a good climate and fertile soil compared with B. On the other hand, B's workers are far more skilful. All factors are fully employed.

When both countries divide their factors equally between the production of wheat and cars, they can produce as follows:

Country	Wheat (units)	Cars (units)
A	500	100
B	100	500
Total production	600	600

But if A specialises in producing wheat and B concentrates on cars, total production would be 1,000 wheat and 1,000 cars. There is thus a net gain of 400 wheat and 400 cars to be shared between them.

Here the gains are obvious, because A has an *absolute* advantage in producing wheat and B in producing cars. But suppose A also has skilled labour and capital, and is better at producing both wheat and cars, as follows:

Country	Wheat (units)	Cars (units)
A	500	300
B	400	100
Total production	900	400
(no specialisation)		

Are there still gains to be achieved by specialisation? S Provided the rate at which cars can be exchanged for wheat lies within the range $3-1\frac{1}{4}$, the answer is yes. The reason for this is that A's superiority in producing cars is far more pronounced than her superiority in producing wheat for with the same factors she can produce three cars for every one of B's, but only one-and-a-quarter units of wheat for one of B's. Comparative, rather than absolute, advantages are what are really important. The result is that if A specialises in producing cars, leaving B to produce wheat, total production will be 600 cars and 800 wheat.

Suppose now that world conditions of demand and supply are such that 2 wheat exchange for 1 car; i.e. the price of cars is exactly twice that of wheat. A now exchanges 200 cars for 400 wheat, giving her a total of 400 wheat and 400 cars, and B 400 wheat and 200 cars.

It can be seen, therefore, that through specialisation B is 100 cars better off. But has specialisation improved A's position? She now has 400 cars but only 400 wheat, a gain of 100 cars but a loss of 100 wheat. But by her own production she would have had to go without $166\frac{2}{3}$ wheat in order to obtain the extra 100 cars. Thus we can conclude that she too is better off.

The above explanation must be refined to allow for:

(1) DEMAND

The law of comparative costs merely shows how two countries can specialise to advantage when their opportunity costs differ. But until we know the demand for goods we cannot say definitely whether specialisation will take place or, if it does, to what extent. Thus, although a country may be favourably placed to produce certain goods, a large home demand and thus a relatively high price may mean that it is a net importer of that good.

(2) TRANSPORT AND TRADING COSTS

These reduce possible gains and therefore make for less specialisation. Indeed, it is conceivable that transport costs could so offset A's superiority in making cars that B found it better to produce her own requirements.

(3) CHANGES IN THE CONDITIONS OF SUPPLY

Few production advantages are permanent. Climate and, to a large extent, mineral deposits persist, but new techniques can make factors more productive. Thus India now exports cotton goods to Britain!

(4) INTERFERENCE BY NATIONS WITH THE FREE MOVEMENT OF GOODS – BY CUSTOMS DUTIES, ETC. (see pp. 287–90)

(5) THE POSSIBILITY OF DIMINISHING RETURNS SETTING IN AS THE PRODUCTION OF A GOOD INCREASES

The theory as stated assumes that, at all stages of production, wheat can always be produced instead of cars by both A and B at a constant ratio. Thus at any output, A can have 5 wheat instead of 3 cars and B 4 wheat instead of 1 car. But it is likely that as B increases her output of wheat, diminishing returns set in, for inferior land and labour have to be used. Thus instead of getting 4 additional wheat for 1 car, she receives only 3, and later only 2, and so on. The same applies, too, as the production of cars is increased by A. Eventually, therefore, it pays to specialise no longer, A can obtain her wheat cheaper by producing herself than by buying it on the world market, and the same applies to B as regards cars. Diminishing returns, and thus increasing costs, usually mean in practice that there is only partial specialisation – up to the point where opportunity costs are less than those offered by the terms of trade, that is, the quantity of imports obtained for a given quantity of exports. Thereafter it is better for a country to produce the good itself.

(ii) *By expanding the market, international trade enables the benefits of large-scale production to be obtained*

Many products, e.g. computers, pharmaceuticals, aircraft and cars, are produced under conditions of decreasing cost. Here the home market is too small to exploit fully the advantages of large-scale production. This applies particularly to small countries such as Switzerland. In such cases international trade lowers costs per unit of output.

(iii) *International trade increases competition and thereby promotes efficiency in production*

As we have seen, any restriction of the market makes it easier for one seller to gain control. In contrast, international trade increases competition. A government must always consider the risk of a monopoly developing when it gives protection to the home industry by tariffs, etc.

(iv) *International trade promotes beneficial political links between countries*

Examples of this are the EC, and with the Commonwealth trade is still an important link.

24.2   A note on the terms of trade and foreign currency exchange rates

We said above that, for trade to take place between two countries, the rate at which goods can be exchanged in the market must be better than the production possibility ratios between the commodities which each country could achieve for itself.

In practice the market terms of trade depend upon: (a) real forces – the conditions of demand and supply for the respective goods; (b) monetary influences – the rate at which one country's currency exchanges for another's. A change in either can affect the actual terms of trade.

(a) Terms of trade

The terms of trade express the relationship between the price of exports and the price of imports. Since these are so varied and because the main interest centres on relative changes in the terms of trade between one period and another, such changes are measured by index numbers for both exports and imports. This involves choosing a base year, and revaluing given 'baskets' for each.

The terms of trade (Table 24.1) is itself an index, equalling the index for exports divided by the index for imports × 100 for any given year.

Thus a fall in the price of country *A*'s exports relative to that of her imports would produce a fall in the terms of trade, and we should say that her terms of trade have 'worsened', 'moved against her' or 'become less favourable', and vice versa.

Table 24.1 shows that, apart from a dip in 1986, the terms of trade for the UK have held fairly steady.

(b) Foreign currency exchange rates

Exchange rates are determined by the demand for and supply of currencies on the foreign exchange market. Worldwide dealing (arbitrage operations) ensure that world exchange rates prevail.

Demand for and supply of a foreign currency depend on: (i) a country's overall balance of payments on current account (see pp. 295): (ii) movements of capital – long-term, short-term and speculative (see pp. 296). The underlying factors which affect trade and capital movements are:

Table 24.1   The terms of trade of the UK, 1984–91 (base year 1985)

Year	Export unit-value index (1)	Import unit-value index (2)	Terms of Trade (1) ÷ (2) × 100
1984	95.0	95.3	99.7
1985	100.0	100.0	100.0
1986	90.1	95.4	94.4
1987	93.5	98.0	95.8
1988	93.4	97.0	96.3
1989	100.8	104.2	96.7
1990	106.2	108.1	98.2
1991	106.4	108.1	98.1

Source: *Monthly Digest of Statistics.*

1. *Relative prices of goods and services traded between the different countries.* Thus, if there is inflation in the UK relative to other countries, the cost in sterling of her exports will rise, and importers will turn to alternative sources. Hence the demand for sterling to pay for UK exports will fall.
2. *Changes in real income.* Thus recession in Germany could decrease her demand for UK car components, and so reduce her demand for sterling.
3. *Long-term capital investment projects.*
4. *Relative rates of interest offered* by countries on both long-term and short-term loans.
5. *Expected future movements in exchange rates.* If, for example, the sterling exchange rate is expected to fall, speculators will use sterling to buy other currencies now, and re-sell for more sterling once it is devalued. Indeed such speculation tends to increase the possibility of devaluation.
6. *The political stability of a country* and international confidence in the soundness of its economic policies.

24.3   The pattern of the UK's overseas trade

(a) Trade with other countries

From our study of the reasons for international trade we can deduce the likely pattern of the UK's trade. Since, in the first place, trade arises because resources are unequally distributed, we have to ask: (i) What are the factors of production of which the UK can be said to have relatively plentiful supply? (ii) What are the factors in which she is deficient?

In answer to (i), we can point to her oil and natural gas reserves, her skilled workers, factories and machinery. In addition she has a high proportion of very skilled and highly educated administrators, engineers and technicians, researchers, and commercial and financial experts. All such persons can render services to

other countries, particularly the less developed. Thus administrators go abroad to start businesses and engineers plan and construct buildings and bridges, while commercial and financial experts and institutions perform services for countries other than the UK.

As regards (ii), however, the UK lacks land; plentiful supplies of very cheap, unskilled labour; certain minerals (such as nickel, zinc, aluminium and copper); certain chemicals (such as sulphur and nitrates); and the climate which is necessary, in terms of both warmth and rainfall, for the production of many foodstuffs (such as cane sugar, vegetable oils and tropical fruit), beverages (such as tea, coffee and cocoa) and raw materials (such as cotton, rubber and tobacco).

(b) The commodities of the UK's international trade

The relative supply of the UK's resources suggests that she will mostly export manufactured goods and also render services to other countries. In actual fact, her main *net* exports in value are beverages (e.g. whisky), chemical materials, medicinal and pharmaceutical products, perfume and toilet materials, iron and steel manufactures, power generating machinery, electrical and electronic equipment, aerospace products, scientific instruments and petroleum. In addition, the UK supplies one-tenth of the world's services.

Similarly she will import food, tropical products, together with raw materials, minerals and chemicals in which she is deficient. Thus we find her main *net* imports in value are meat, animal feeding stuffs, vegetables and fruit, coffee, tea, cocoa and spices, wood and cork, paper and pulp metalliferous ores and scrap, explosives, textiles, clothing, footwear, electrical appliances, photographic equipment and road vehicles.

Table 24.2   The UK's imports (cif) and exports (fob) 1991 (by value)

	Imports		Exports	
	£m	%	£m	%
Food, beverages and tobacco	12 325	10.4	7 748	7.4
Crude materials	4 678	3.9	1 920	1.9
Mineral fuels and lubricants	7 511	6.3	7 169	6.8
Manufactured goods	9 160	77.6	86 137	82.1
Miscellaneous (postat packages; animal and vegetable oils; commodities not classified elsewhere)	2 112	1.8	1 903	1.8
	118 786	100	104 877	100

Source: *Annual Abstract of Statistics.*
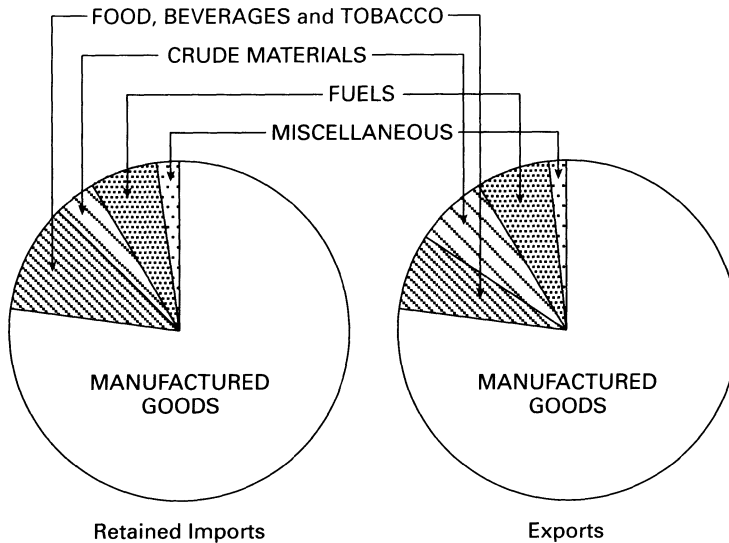
Figure 24.1   Percentage distribution of the UK's imports and exports, 1991

But the broad divisions of exports and imports (Figure 24.1) requires some explanation. Food is forming a decreasing proportion of UK imports, partly because increased home production now supplies two-thirds of her requirements, and partly because spending on food takes a smaller percentage of an increasing income.

What may appear more remarkable, however, is the preponderance of trade in manufactured goods as regards both exports (82%) and imports (78%). Indeed such trade is increasing in importance. The main reason is that specialisation is no longer confined to either manufacturing or to agriculture. Now, even within manufacturing, different countries concentrate on producing particular goods or even components. Thus, in cars the Mercedes appeals to certain people in the UK, while the Rolls Royce and Jaguar may be wanted in Germany. Such specialisation can give rise to considerable trade between countries which have reached the same stage of industrial development (see also Figure 24.2).

A further reason is that one-third of manufactured goods consists of semi-manufactures because basic material producers are now adding value by processing themselves e.g. in chemicals, textiles, iron and steel, and non-ferrous metals. Indeed since 1983 the UK has been net importer in value of manufactured goods (see also pp. 318–20).

(c) The countries with which the UK trades

Again an analysis of the UK's resources compared with those of other countries suggests that she will import goods from countries having relatively much agricultural land, enjoying a tropical or semi-tropical climate or possessing the minerals and chemicals which she herself lacks. Where these countries need the UK's manufactured goods, as in the case of Norway, imports from them can be

WHERE THE UK BUYS		WHERE THE UK SELLS
17718	Germany	14035
13093	USA	11431
11075	France	11597
9969	Netherlands	8258
6754	Japan	2260
6379	Italy	6410
5472	Belgium & Luxembourg	5873
4416	Irish Republic	5295
4248	Switzerland	2105
4163	Norway	1358
3142	Sweden	2471
2628	Spain	4279
2227	Denmark	1409
1972	Hong Kong	1387
1923	Canada	1701
963	Saudi Arabia	2255
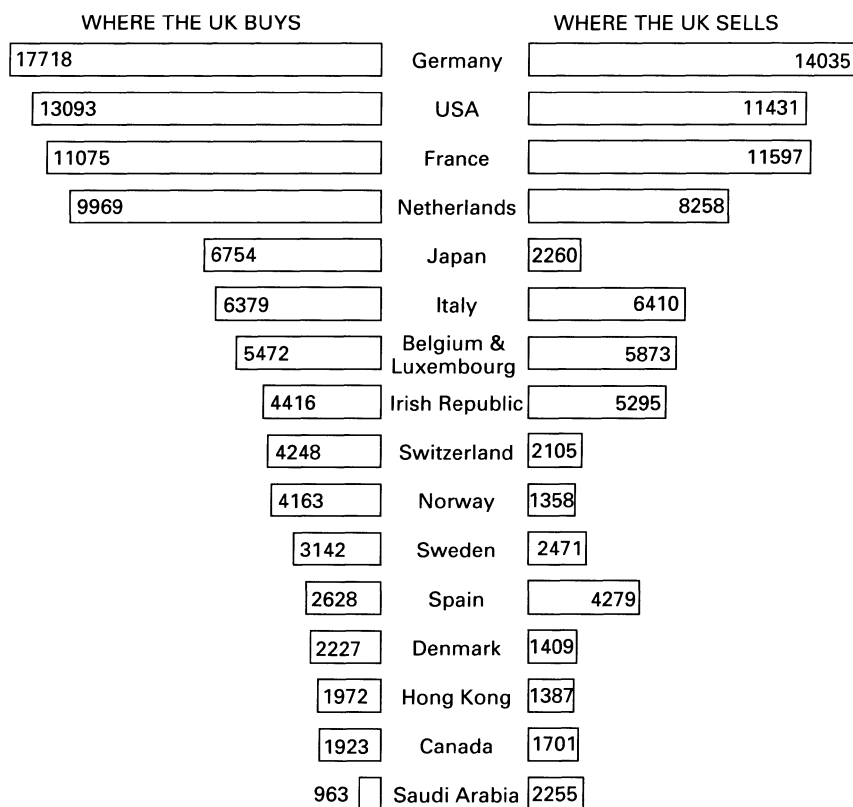
Figure 24.2 The principal exporters to and importers from the UK, 1991 (£m)

paid for directly by the export of such goods as timber. But where the country, like the USA, does not require manufactured goods, imports have to be paid for indirectly. This is achieved through triangular or multilateral trade. Thus Malaysia exports tin to the USA, but the latter sends comparatively little to them directly in exchange. Instead she exports such goods as cotton, tobacco and machinery to the UK, which in return settles the USA's bill from Malaysia by sending Malaysia manufactured goods.

(d) The UK's trade with the EC
In 1972 the UK's trade with the other eleven members of the EC was just over a third. Table 24.3 shows that in 1991 it is over one-half, and is increasing especially with the Single Market.

Figure 24.2 shows that six out of the UK's eight main markets in both imports and exports were EC countries.

Table 24.3 Pecentage distribution of the UK's trade between the EC and the rest of the world, 1982 and 1991

	Imports %		Exports %	
	1982	*1991*	*1982*	*1991*
EC	44.3	51.6	41.6	56.5
Rest of the world	55.7	48.4	58.4	43.5
TOTAL	100	100	100	100

Source: *Annual Abstract of Statistics.*

24.4   Free trade and protection

(a) Controlling international trade

Our earlier analysis suggests that trade should be as free as possible, for only then can maximum specialisation according to the law of comparative advantage take place. In practice, however, all countries follow policies which, to varying degrees, prevent goods from moving freely in response to differences in relative prices. Methods vary.

(i) *Customs duties*

Custom duties, e.g. the common external tariffs of the Common Market, are both revenue-raising and protective. They become protective when the imported good bears a higher rate of tax than the similar home-produced good.

(ii) *Subsidies*

While countries which subscribe to the General Agreement on Tariffs and Trade (GATT) cannot follow a policy of 'dumping' exports by giving direct subsidies, the volume and pattern of international trade may be influenced indirectly by other means, e.g. government assistance to the shipbuilding industry. Less obviously, welfare benefits, e.g. child-benefits and income supplements which keep down labour costs, may give one country a price advantage over another.

(iii) *Quotas*

If demand is inelastic, the increase in price resulting from a customs duty will have little effect on the quantity imported. Thus, to restrict imports of a good to a definite quantity, quotas must be imposed.

Compared with duties, quotas have two main disadvantages. (1) As a result of the artificial shortage of supply, the price may be increased by the foreign supplier or by the importer. Hence, unless the government also introduces price control, they gain at the expense of consumers. (2) Quotas are usually based on a

firm's past imports, which makes the economy rigid by penalising the efficient firm wishing to expand.

To avoid having formal quotas imposed, 'voluntary export restraints' may be agreed, e.g. on the import of Japanese cars.

(iv) *Exchange control*
A tighter check on the amount spent on imported goods can be achieved if quotas are fixed in terms of foreign currency. This necessitates some form of exchange control. All earnings of foreign currency and claims to foreign currency have to be handed over to the government and goods can be imported only under licence. Thus the government, not the free market, decides the priorities for imports.

(v) *Physical controls*
A complete ban – an embargo – may be placed on the import or export of certain goods. Thus narcotics cannot be imported, while the export of some strategic goods to certain countries is forbidden. Similarly, imposing strict technical standards for certain goods (e.g. milk) and quarantine and health regulations (e.g. dogs and parrots) make trade in them more difficult.

(b) **Reasons for government control of international trade**
In general trade is controlled because governments think nationally rather than internationally. Although people as a whole lose when trade is restricted, those of a particular country may gain.

Many reasons are put forward to justify controls. Occasionally they have some logical justification; more usually they stem from sectional interests seeking to gain advantages. We can, therefore, examine the arguments under three main headings: (i) those based on strategic, political, social and moral grounds; (ii) those having some economic basis; and (iii) those depending on shallow economic thinking.

(i) *Non-economic arguments*

(1) TO ENCOURAGE THE PRODUCT OF A GOOD OF STRATEGIC IMPORTANCE
Where a nation is dependent on another for a good strategic importance, there is a danger of its supply being cut off in the event of war. Thus one argument for sub-sidising aircraft production in the UK is that it will ensure the survival of technical know-how, plant and skilled labour.

(2) TO FOSTER CLOSER POLITICAL TIES
As a member of the EC, Britain must impose a common external tariff as part of a movement towards political as well as economic unity.

(3) TO PROSECUTE POLITICAL OBJECTIVES
Trade can be a weapon of foreign policy, e.g. in the sanctions against Serbia following the invasion of Bosnia.

(4) TO PROMOTE SOCIAL POLICIES
Although in the past Britain has subsidised her agriculture mainly for strategic reasons, today the purposes are basically social – to avoid depression in rural districts.

(ii) *Economic arguments having some justification*

(1) TO RAISE REVENUE(see p. 268)

(2) TO IMPROVE THE TERMS OF TRADE
The incidence of a selective tax is shared between producer and consumer according to the relative elasticities of supply and demand (see p. 273). A government can therefore levy a tax on an imported good to improve the terms of trade if demand for the good is more elastic than the supply, for the increase in price is borne mainly by the producer, while the government has the proceeds of the tax. In practice this requires that: (a) the producing country has no alternative markets to which supplies can be easily diverted; (b) her factors of production have few alternative uses; and (c) the demand for the exports of the country imposing the tariff is unaffected by the loss of income suffered by countries who now find their sales abroad reduced.

(3) TO PROTECT AN 'INFANT INDUSTRY'
It may be possible to establish an industry if during its infancy it is given protection from well-established foreign competitors already producing on a large scale. It is argued that eventually the 'infant' will be strong enough to compete successfully. Britain's car industry, for instance, initially benefited from such protection.

In practice, industries tend to rely on this protection, so that tariffs are never withdrawn; for example, American duties on manufactured goods imposed in the eighteenth century still persist today. Moreover, industries are often encouraged which without protection would have no chance of survival. This leads to maldistribution of a country's resources.

(4) TO ENABLE AN INDUSTRY TO DECLINE GRADUALLY
Fundamental changes in demand for a good may severely hit an industry. Such, for instance, was the fate of the British cotton industry in the 1970s. Restrictions on imports can cushion the shock, giving the industry more time to contract or restructure.

(5) TO PREVENT DUMPING
Goods may be sold abroad at a lower price than in the home market. This may be possible because: (a) producers are given export subsidies; (b) price discrimination by a monopoly is possible; or (c) it enables the producer to obtain the advantages of decreasing costs. People in the importing country benefit directly from the lower prices. If, however, the exporter is trying to establish a monopoly

which can be exploited once home producers have been driven out, there is a case
for protecting the home market.

(6)   TO CORRECT A TEMPORARY BALANCE OF PAYMENTS DIS-EQUILIBRIUM

(iii) *Economic arguments having little validity*

(1)   TO RETALIATE AGAINST THE TARIFFS OF ANOTHER COUNTRY
The threat of a retaliatory tariff may be used to influence another country to
modify a restrictive policy. While this may be successful, it can induce counter-
retaliation, with everybody losing.

(2)   TO MAINTAIN HOME EMPLOYMENT IN A PERIOD OF DEPRESSION
Countries may place restrictions on imports to promote employment in the manu-
facture of home-produced goods. The difficulty is that other countries retaliate,
thereby leading to an all-round contraction in world trade. GATT was set up to
prevent this from happening (see p. 291).

(3) TO PROTECT HOME INDUSTRIES FROM 'UNFAIR' FOREIGN COMPETITION
The demand that British workers must be protected from competition by cheap,
'sweated' foreign labour usually comes from the industry facing competition.
The argument, however, has little economic justification. First, it runs counter to
the principle that a country should specialise where it has the greatest advantage.
That advantage may be cheap labour. Second, low wages do not necessarily
denote low labour costs. Wages may be low because labour is inefficient through
low productivity. What is really significant is the wage cost per unit of output.
Thus the USA can export manufactured goods to the UK even though her labour
is the most highly paid in the world. The threatened industry can compete by
improving productivity to reduce wage-cost per unit. Third, a tax on the goods of
a poor country merely makes the country poorer and its labour cheaper. The way
to raise wages (and the price of the good produced) is to increase demand in for-
eign markets. Indeed, if imports from poor countries are restricted, other help has
to be given. They prefer 'trade to aid'. Fourth, they have less to spend on
Britain's exports. Fifth, the policy may lead to retaliation or aggressive competi-
tion elsewhere, thereby making it more difficult for the protecting country to sell
abroad. One reason why Japan captured many of Britain's foreign markets for
cotton goods was that her sales to Britain were restricted by protective barriers.
Last, restrictions on competitive imports may allow home firms to raise their
prices. If wage increases result, exports of goods generally could fall through
higher prices.
   While restriction of trade tends to lower living standards, there may be benefits
– economic, political and social. Thus protection may be given to an industry
because home workers cannot adjust quickly to other occupations or industries.
Usually, however, such economic gains are doubtful. Others cannot be measured,
and it has to be left to politicians to decide where the balance of advantage lies. It
must, however, always be remembered that protection creates vested interests
opposed to subsequent removal.

(c) The General Agreement on Tariffs and Trade (GATT)
The General Agreement on Tariffs and Trade, established in 1947, has three
major objectives: (i) to reduce existing trade barriers; (ii) to eliminate dis-
crimination in international trade; and (iii) to prevent the establishment of further
trade barriers by getting nations to agree to consult one another rather than take
unilateral action. It operates as follows.

Member nations meet periodically to 'agree' on a round of tariff reductions.
Here the 'most-favoured-nation' principle applies – any tariff concession granted
by one country to another must automatically apply to all other participating
countries. Thus, if the EC agrees to reduce her tariff on American automatic
vending machines by 5% in exchange for a 5% reduction in the American tariff
on EC man-made fibres, then both concessions must be extended to every other
member of GATT.

Today over one hundred nations subscribe to GATT. Through the organisa-
tion, a progressive reduction in tariffs has been achieved, and the principle
has been established that problems of international trade should be settled by
co-operative discussion rather than by independent unilateral action. But dif-
ficulties have arisen. (i) The principle of reciprocity means that low-tariff
countries have to begin from inferior bargaining positions, and the con-
cessions they can make are thus limited. Such countries may, therefore, pre-
fer a low-tariff regional arrangement, such as the EC. (ii) In certain
circumstances the 'most-favoured-nation' principle may deter a country from
making a tariff reduction to another country for the simple reason that it has
to be applied to all. (iii) The articles of the agreement have had to be waived
to allow for special circumstances – balance-of-payments difficulties, pro-
tection of her agriculture, the establishment of 'infant' industries in less-
developed countries and the discriminatory character of the EC. (iv) While
the GATT has been successful in dealing with tariffs and many physical bar-
riers, it has been by-passed by the new forms of protection – voluntary agree-
ment restraints, orderly marketing arrangements, subsidies for special groups
of exports, and trading requirements as conditions for overseas investment.
(v) the GATT rules will eventually have to be extended to cover *services,*
which now account for a quarter of world trade, and *intellectual property
rights* (patents, copyrights and trademarks).

The *Uruguay Round*, which started in 1986, took seven years to complete,
largely because the USA required reform of the EC's Common Agricultural
Policy to prevent subsidised surplus produce depressing world prices.

Although agreement was reached on liberalising trade in services, the less-
developed countries would not concede on intellectual property rights.

24.5   The balance of payments

(a) Paying for imports
Occasionally international trade takes the form of a barter arrangement, one
country agreeing to take so much of another country's produce in exchange for so
much of its own. Normally, however, exchanges are arranged by private traders

who, according to relative prices, decide whether it is profitable to export and import goods.

But each country has its own currency. This fact is important for two reasons: (i) sufficient foreign currency has to be obtained to pay for imports; and (ii) a rate has to be established at which one currency will exchange for another.

We can approach the question of how imports are paid for by considering the purchases made by a housewife, Mrs Jones. Each week she buys a variety of goods. However, there are at least seven sources from which she could obtain the money to pay for them.

The most usual is the week's earnings. From her husband's allowance, say, Mrs Jones pays the shopkeeper as she collects her goods. But what Mrs Jones is really doing is exchanging the goods which Mr Jones has specialised in producing for the other goods needed. Thus, if Mr Jones is a tailor, the money from the suits he sells buys Mrs Jones the goods she needs. Furthermore, money is often earned not by making goods but by performing a service. Thus Mrs Jones herself may earn wages by working for the shopkeeper. Last interest on savings may provide some current income. Provided that all the weekly expenses are met out of this combined weekly income, we should say that the Jones family was 'paying its way'.

It might happen, however, that Mrs Jones's expenditure was not covered by the current weekly income. This might occur, for instance, because she bought a costly good, such as a dish-washer. In such circumstances, she would have to raise the money from other sources. First, she could draw on her savings. Second, she could sell household goods, such as the piano or the TV set, for which she had a less urgent need. Third, she might be able to borrow the money from a friend or, what amounts to the same thing, ask the shopkeeper to forgo being paid for the time being. Finally, if she were extremely fortunate, she might be able to obtain a gift of money, say from a doting father. Such methods of payment would be fairly satisfactory for a good which is in use over a long period, provided that her savings were gradually replenished, or the assets sold were replaced by others of equal value, or the loan was repaid during the lifetime of the good. Otherwise Mrs Jones would not be paying her way. If over-spending continued, her savings would eventually run out, her home would be sold up, and she would be unable to obtain any more loans or credit from the shopkeeper.

Broadly speaking, a nation trading with other nations is in exactly the same position. The same alternatives are open to it in paying for imports. The main source is receipts from current exports. Figure 24.3 shows how exports earn foreign currency. Importing and exporting are arranged by firms, and payments are arranged through banks, who exchange the currency of one country for the currency of another *provided that they have the necessary reserves of that currency.* Such reserves are earned by customers who export to foreign countries.

Let us assume that £1 sterling exchanges for $1.50 and that there are no currency restrictions. Suppose a British merchant, X, wishes to import cotton from A in the USA to the value of £100,000. The American exporter requires payment in dollars, for all his payments, e.g. his workers' wages, have to be made in dollars. Hence the importer goes to his bank, pays in £100,000 and arranges a 'documen-
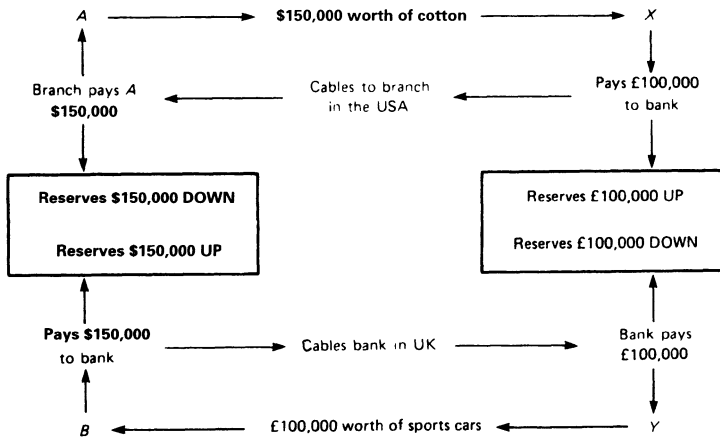
Figure 24.3 How exports pay for imports

tary credit'. The bank cables its branch in New York, authorising it to make the equivalent dollar payment to *A* on production of the necessary documents, e.g. the bill of lading (see p. 179). (Most banks have branches in foreign capitals; if not, they engage local banks to act for them.) But how is it that the branch has dollars available to honour the draft?

We can see the answer if we imagine that another British firm, *Y*, has sold £100,000 worth of sports car to an importer, *B*, in the USA. This firm wants payment in £ sterling. Hence the American importer of the car pays $150,000 into his bank in the USA, and the same procedure follows. It is obvious that the two transactions – buying cotton from the USA and selling sports cars to the USA – balance each other. The British bank's branch has had to pay out dollars, the American sterling. The British bank has received sterling, the American bank dollars. If the two get together, their requirements match. (In practice it is more likely that they would meet their needs through the foreign-exchange market.) The dollars to pay for the cotton are obtained by selling the sports cars, and *vice versa*. In short, exports pay for imports.

(b) 'Exports' in the wider sense
In this connection the term 'exports' needs qualification. In the same way that Mrs Jones received payment for her services to the shopkeeper, so a nation may receive payment not only for goods but also for services rendered to other countries. Goods exported are termed 'visible exports' because they can be seen and recorded as they cross frontiers. But services performed for people of other countries cannot be seen and recorded; they are therefore called 'invisible exports'. Nevertheless, since services involve payments by persons from abroad, they are exports.

The main sources of invisible earnings and payments are:
(i) *government expenditure abroad,* e.g. on overseas garrisons and diplomatic services, overseas aid.

(ii) *shipping services*, e.g. for an American travelling in the QE2 or shipping exports in a British merchantman;

(iii) *civil aviation*;

(iv) *travel*, e.g. sterling required by an American tourist for spending on a visit to London;

(v) *other services*, e.g. royalties earned on books and records, and income from the transactions of overseas oil companies which ship direct from wells and refineries abroad to other countries;

(vi) *interest, profits and dividends from overseas investments*;

(vii) *private transfers*, e.g. remittances to relatives abroad.

Payments for any of the above transactions involve changing into another country's currency. Thus they represent imports to the paying country and exports to the receiving country.

(c) The balance-of-payments accounts

The accounts presented by a country of its monetary transactions with the rest of the world are known as the 'balance of payments' (Table 24.4).

(i) *Current account*

The current account shows on the one hand the foreign currency which has been *spent* on *imported goods* and *invisibles* in the course of the year, and on the other the foreign currency which has been *earned* by *exporting goods* and *invisibles*.

That part of the current account showing payments for just the *goods* exported and imported is known as the *visible balance* (formerly the *balance of trade*). Where the value of goods exported exceeds the value of goods imported, we say that there is a favourable visible balance. If the opposite occurs, the visible balance is unfavourable. Too much, however, must not be read into the terms 'favourable' and 'unfavourable'. In the first place we have to know the reasons for the unfavourable balance. It may be brought about, for instance, by an increased demand for raw materials and components as a country moves out of a recession. These will later be exported as manufactured goods. Or a less developed countries may receive more aid, enabling her to import capital goods. Whereas the value of these is shown as 'imports', 'aid' will appear in the Financial Account. Secondly, a favourable or unfavourable visible balance can be reversed when the invisibles are taken into account.

When we add to the visible balance, payments and income on the invisible items, we have what is known as the *current balance*.

There is no special reason why earnings from goods and invisibles exported between 1 January and 31 December in any one year should equal expenditure on the goods and invisibles imported during that period. In fact, it would be an extraordinary coincidence if they did so. How often does what you earn during the week tally *exactly* with what you spend?

The current account is therefore likely to show a difference between earnings and expenditure. When the *value* of goods and invisibles exported exceeds the *value* of goods and invisibles imported, we say that there is a surplus current balance; when the reverse occurs, we say that there is a deficit current balance.

Table 24.4   The Balance of Payments of the UK, 1991 (£m)

CURRENT ACCOUNT
  Visible trade

Exports (fob)	+103 413	
Imports (fob)	−113 703	
Visible balance		−10 290
Invisibles (net)		
Government	−2 396	
Sea transport	+15	
Civil aviation	−470	
Travel	−2 660	
Financial and other services	+105 501	
Interest, profits and dividends	+329	
Private and government transfers	−1 349	
Invisible balance		+3 969
CURRENT BALANCE		−6 321

FINANCIAL ACCOUNT
  Transactions in external assets and liabilities:

Investment abroad (net)	−12 497	
Borrowing abroad (net)	+20 408	
Official reserves, addition to	−2 662	
TOTAL	+5 249	
Balancing item	+1 072	
CAPITAL BALANCE		+6 321

Source: *Annual Abstract of Statistics.*

The importance of the current balance is that it shows how far a country is paying its way.

However, the current account is only part of the statement covering a nation's overseas financial transactions. Capital flows must also be scrutinised. As we shall see, a current deficit need cause no alarm if it is covered by borrowing which will be put to a productive use. On the other hand, a current surplus may be insufficient to offset a heavy drain on the reserves through the outward movement of short- and long-term capital. The balance-of-payments statement must be examined as a whole.

(ii) *The financial account*
The *financial account* sets out the currency flow generated by current account balances and capital movements.

If the current-account transactions were a country's only dealings with the world, the balance-of-payments accounts would be quite simple. A surplus of £100 million, for example, would add that amount to the reserves or allow the country to invest that amount overseas or to pay off short-term borrowings from the International Monetary Fund (IMF) or other foreign creditors. A deficit of £100 million would reduce the reserves by that amount or have to be financed by disinvestment or short-term borrowing abroad.

But *capital* flows also affect a country's ability to build up reserves or to pay off debts. Thus investment by private persons resident in the UK in factories or plant overseas (whether directly or by the purchase of shares), or a loan by the British government to a less-developed country, leads to an outflow of capital and the spending of foreign currency (negative sign in the accounts). Similarly, investment in the UK by persons overseas, or borrowing abroad by the British government, local authorities, or companies, leads to an inflow of foreign capital and the receipt of foreign currency (+ sign).

Whereas the current account covers *income* earning and spending in the course of the year, 'transactions in external assets and liabilities' show the movement of *capital* in and out of the country. This capital may be short- or long-term.

Short-term capital movements arise from the transfer of liquid funds to and from Britain. Because London is a world financial market centre, foreigners hold bank balances or short-term bills there. These short-term funds can move quickly from country to country to take advantage of higher interest rates or to guard against an exchange rate depreciation. They are thus often referred to as 'hot money'.

Long-term capital investment by British residents in factories or plant overseas (whether directly or by the purchase of shares), or a loan by the British government (e.g. to a less developed country or an international institution) leads to an outflow of capital. Similarly, investment in the UK by persons overseas or borrowing from abroad by the British government, local authorities, or companies leads to an inflow of foreign capital.

Any movement of capital out of Britain gives rise to a demand for foreign currency, a movement into Britain from abroad leads to the receipt of foreign currency.

No distinction is made between short- and long-term investment in presenting the overall balance of payments. In fact much of Britain's overseas investment is financed by short-term capital borrowed from foreigners, e.g. from the pool of Eurocurrency deposited in London. To the extent that this occurs, there is no net outflow of foreign currency. Britain's overseas investment which is undertaken in order to make a profit is, in fact, like private business ventures. And, just as the shopkeeper borrows from the bank to cover the holding of stocks before Christmas, so the UK borrows to finance investment overseas in factories, plantations, oil wells, nickel-mines, etc.

Thus the UK's balance-of-payments accounts concentrate on what is really significant to Britain – the extent to which currency flows as a whole influence the £ sterling exchange rate and her reserves of gold and foreign currencies.

The balancing item arises as follows. When all recorded capital transactions are added to the current balance, the total never adds up exactly to the amount of foreign currency the country has in fact gained or lost, which is known precisely to the Bank of England. Government spending overseas, for instance, is easier to record exactly than the foreign spending of people taking holidays abroad. Exports, too, may go abroad in December, but payments from them come in the following February.

A 'balancing item' is therefore added to make up the difference between the total value of the transactions recorded and the precise accounts kept by the Bank of England. If the balancing item is '+', it means that more foreign currency has actually come in than the estimates of transactions have indicated. When there is a '−' balance item, the opposite is the case.

(d) An examination of the UK's balance of payments, 1991

The above explanation can be illustrated by examining the UK's Balance of Payments for 1991 (Figure 24.4).

Imports exceeded exports; there was a visible balance deficit of £10290 m. On invisibles, the UK had a favourable balance of £3969 m. The overall deficit on the Current Balance was thus £6321 m.
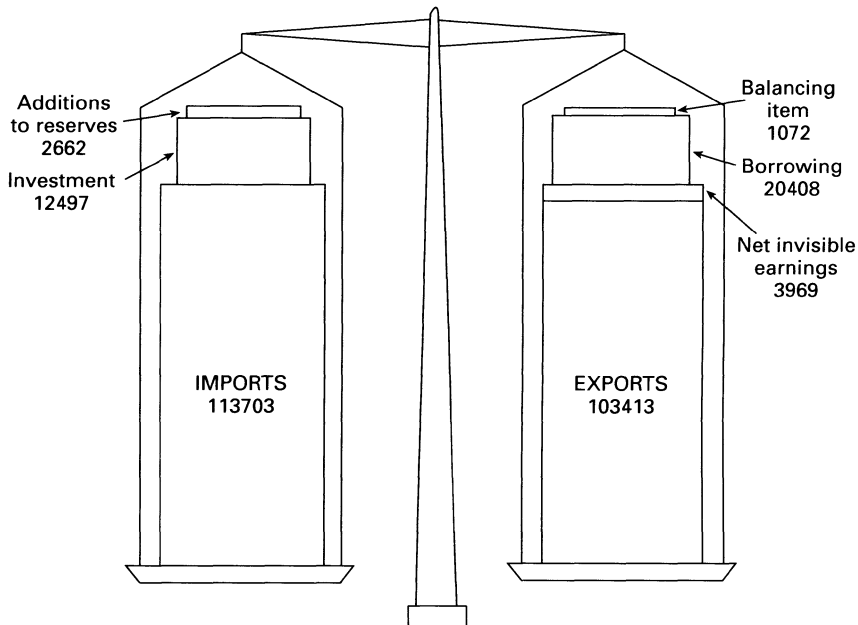
Figure 24.4 The Balance of Payments, 1991

In addition to the current account deficit of £6321 m, net investment of £12497 m abroad by Britain had to be covered. This was achieved by net borrowing of £20408 m, together with a balancing item of £1072 m, which allowed for an addition to the reserves of £2662 m.

24.6   Balance-of-payments problems

(a) A persistent deficit on the current account

It may be that a country has a deficit in a certain year because, for instance, it spent heavily on imported stocks of raw materials. The following year, however, those materials may be made into manufactured goods which are exported, and the deficit thus turned into a favourable balance. If so, there is little need to worry about a deficit, for it can be financed from the reserves or from short-term borrowing.

A different situation occurs where a deficit continues from year to year. The reserves could run out, and creditors will not lend indefinitely. Action has to be taken, therefore, to remedy the situation.

A first-aid measure is for the monetary authorities to take action to raise short-term interest rates in order to attract foreign loans. Eventually, however, exports must be increased in value, and imports decreased.

Two basic policies can be followed: (i) reducing expenditure on imports; and (ii) switching expenditure, so that foreigners spend more on British exports, and consumers at home spend less on imports in favour of home-produced goods. Both policies can be followed simultaneously (though with a different emphasis on each), but to clarify the issues it is better to consider them separately.

(i) *Reducing expenditure on imports: deflation*

Expenditure on imports may be reduced forcibly by the government's imposing import duties, quotas and exchange controls. However, not only is such a policy likely to arouse hostility and even to lead to retaliation by foreign countries, but it offends the spirit of GATT and of the IMF. However, as we saw in chapter 14, imports increase as income expands. Thus one way in which the value of imports can be brought into line with that of exports is by reducing income.

Figure 24.5 explains the situation. To simplify we assume an economy with no government spending or taxation, injections consisting of autonomous investment and exports, while leaks of saving and imports are related to income. At the current level of income, Y, there is a current balance-of-payments deficit, DF. Assuming exports are maintained, this deficit can be elimated by bringing down the level of investment to I', reducing income to Y'.

Such a deflationary policy would also tend to put a brake on any rise in home prices. More important, it allows adjustment to take place without altering currency-exchange rates (see below). This has the advantage that it facilitates international trade by removing the uncertainty associated with fluctuating exchange rates when negotiating long-term contracts or making loans.

But there are serious disadvantages. Unless home prices are flexible downwards, a deflationary policy only succeeds at the expense of creating unemploy-
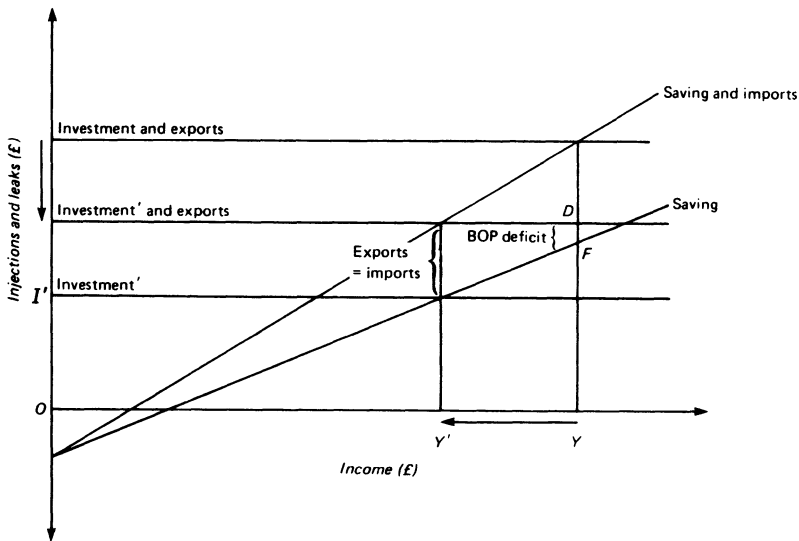
Figure 24.5 Achieving balance-of-payments equilibrium by deflation

ment. Moreover the impact on expenditure on imports is less where these have a low income-elasticity of demand, e.g. essential raw materials and components. Finally a reduction of imports to Britain represents a loss of exports by other countries. Deflationary effects on their economies may result in a reduction in their demand for imports, thereby reducing British exports. In short, deflation is a 'beggar-my-neighbour' policy where the benefit is uncertain even to the deflating country.

(ii) Increasing exports and decreasing imports by expenditure switching
Whilst some switching can be enforced by government control on import expenditure, the most effective method is to alter the relative prices of imports and home-produced goods. Prices of internationally traded goods are composed of (i) the home producer's price and (ii) the exchange rate. Thus British exports can be made more competitive by lowering the rate at which the £ sterling exchanges for foreign currencies. Because fewer units of foreign currency have now to be given up to obtain a pound sterling, foreigners can buy British exports more cheaply. If their elasticity of demand is greater than I, more revenue will be earned from exports (provided, of course, that supply by British firms can respond to the extra demand). Similarly, imports to Britain now cost more, encouraging people to switch to the relatively cheaper home-produced goods.

Such corrective exchange-rate variations occur automatically through the forces of demand and supply in a freely operating foreign-exchange market. Thus, if country *A*'s exports to *B* are greater in value than her imports, the demand for *A*'s currency will increase, for importers in *B* will be wanting it to pay for those imports. Consequently the exchange rate will move in favour of *A*. Similarly, if *A*'s imports are greater in value than her exports, the exchange rate will

move against her. In each case, the movement tends to bring about equality in the values of total imports and total exports (see also p. 316).

While the system of floating exchange rates avoids deflation of incomes and thus increased unemployment at home, it has disadvantages. Apart from the fact that it introduces uncertainty into international trade, the prices of imported food-stuffs and raw materials are raised for the depreciating country, thereby adding to the cost of living and leading to claims for wage increases.

Thus the UK has, in the past, adopted different arrangements aimed at stabilis-ing the £ sterling exchange rate – adhering to a 'gold standard', agreeing to a system of 'managed flexibility' according with International Monetary Fund rules and, more recently, by joining the EC's Exchange Rate Mechanism (ERM).

But in each case deflationary pressures have led to a reversion to the floating of the £ sterling on the foreign exchange market (see also p. 317).

(b) The Exchange Equalisation Account
Even with fluctuating exchange rates it is necessary to lessen the effect of speculative short-term capital movements.

The Exchange Equalisation Account operates by the simple application of the laws of price. It has a stock of gold and foreign currencies (mostly borrowed against Treasury bills), and this stock is either replenished or offered on the market according to whether short-term capital is moving into or out of London. For instance, a movement of capital into London from the USA would increase the demand for the £ sterling and drive up its price; the Account can prevent this rise by offering pounds in exchange for dollars. On the other hand, if there were a movement of capital out of London, the Account would offer dollars in exchange for pounds, thereby reducing its stock of dollars and increasing its holding of pounds.

The knowledge that such an Account exists to even out exchange fluctuations has done much to prevent speculation in the value of the pound.

⟨25⟩ The European Community

25.1 Background to the European Community (EC)

(a) Supra-national organisations

The two world wars convinced statesmen in western European countries that some form of political unity was desirable, and in1949 the Council of Europe was created – the basis, it was hoped, of a European parliament. But organisations with definite functions – the Organisation for European Economic Co-operative (founded in 1948), the North Atlantic Treaty Organisation (1949) and the Western European Union (1954) proved more fruitful than did the Council of Europe with its broad aims.

Although these organisations involved co-operation, they were merely voluntary associations, not federal bodies exercising supra-national powers. Although federation was the ultimate aim of European statesmen, they realised that it could only proceed piecemeal and on a functional basis. The first supra-national organisation, the European Coal and Steel Community (ECSC), was formed in 1951 to control the whole of the iron, steel and coal resources of the six member countries – France, West Germany, Italy, Holland, Belgium and Luxembourg.

The success of the ECSC led to the setting up in 1957 of the Atomic Energy Community (EURATOM), a similar organisation, for the peaceful use of atomic energy, and the European Economic Community (EEC), an organisation to develop a 'common market' between the six member countries. All three communities have now been brought within the EC.

(b) Present membership and proposed expansion

When first offered membership of these organisations, Britain refused to join. Instead, with six other nations, she joined the looser European Free Trade Area (EFTA).

But contrary to Britain's expectations, the EEC grew in strength, for difficulties were resolved as they arose. Moreover, Britain's trade with EEC countries increased at a faster rate than that with EFTA, since her goods were more complementary to their economies. Accordingly, after protracted negotiations, the UK joined in 1973. The other members are now: France, Germany, Italy, Belgium, the Netherlands, Luxembourg, Denmark, the Irish Republic, Greece, Spain and Portugal. It is proposed to bring certain EFTA countries – Austria, Finland, Iceland, Norway and Sweden – within a bigger single market by creating a European Economic Area.

25.2   The institutions of the EC

The essential point to grasp is that the 1957 Treaty of Rome set up a 'Community' with its own form of government and institutions.

There are four main institutions:

1. *The Commission*
This is the most important organ of the EC. Its seventeen members (two from the UK) serve for four years. Once chosen, however, the members of the Commission act as an independent body in the interests of the Community as a whole, and not as representatives of the individual governments that have nominated them. Each commissioner is responsible for a separate area of policy.

The Commission is responsible for formulating policy proposals and legislation, promoting the Community interest, trying to reconcile national viewpoints, implementing Council decisions, and supervising the day-to-day running of community policies. As the guardian of the Treaty, it can also initiate action against member states which do not comply with EC rules.

2. *The Council of Ministers*
Each member country sends a cabinet minister (usually according to the subject under discussion) to the Council of Ministers. This is the supreme decision-making body, the Community's 'cabinet'. Its task is to harmonise the Commission's draft Community policies with the wishes of member governments. The Commission's representative in the Council is present by right, but only to discuss, not to vote. Proposals and compromise plans are exchanged between the Council and the Commission. If the Council becomes deadlocked, the Commission reconsiders the proposal in order to accommodate the views of the opposing countries. Over time, 'specialist' Councils have evolved dealing with particular areas of policy, e.g. agriculture, finance, industry, environment.

By the Single European Act of 1987 most single market measures are subject to majority voting with each member's vote weighted roughly according to its population. However, unanimity is still needed for the politically sensitive areas of taxation, the free movement of people, and workers' rights.

Council meetings are chaired by the member state holding the *Presidency,* which rotates every six months. This carries with it the management of Community business, acting particularly as broker in promoting agreement.

The outgoing President also hosts the *European Council*, a summit of heads of government which meets twice a year and sets the agenda for the incoming Presidency. The major problems confronting them are reviewed in an informal and pragmatic way. The object is to suggest loosely-defined strategies so that each member can take into account the impact of its own policies on the others.

3. *The European Parliament*
This is the body of 567 elected members (87 from the UK) directly-elected for five years. Members sit according to party affiliation, not nationality. The Assem-

bly debates Community policies and examines the Community's budget. It can dismiss the Commission by a two-thirds majority.

Its powers were strengthened by the 'cooperation procedure' provided for by the Single Market Act 1987. On most single market proposals, Parliament gives a first opinion when the Commission make a proposal, and then gives a second opinion after the Council has reached a decision in principle.

Community legislation therefore, results from a complex and often lengthy process of consultation and negotiation between the institutions.

4. *The Court of Justice*

This consists of thirteen judges, a President and one from each member country, appointed for a six-year term. Its task is to rule on the interpretation of the Rome Treaty and to adjudicate on complaints, whether from member states, private enterprises or the institutions themselves. Its rulings are binding on member countries, community institutions, and individuals, and have primacy over national law.

5. *Special institutions*

Apart from the four main institutions above, there are also special institutions to deal with particular policies, e.g. the Economic and Social Committee, the European Investment Bank.

25.3   Economic objectives of the EC

The overriding aim of the EC is to integrate the policies of its member countries. Its economic policy is based on two main principles: (1) a customs union, and (2) a common market.

1. *A customs union*

We have to distinguish between a free-trade area and a customs union. The former simply removes tariff barriers between member countries but allows individual members to impose their own rates of duty against outsiders. A customs union goes further. While it too has internal free trade, it also imposes common external tariffs.

The EC has a customs union, since this is essential for an integrated common market. Otherwise goods would enter the market through low-duty countries and be resold in those imposing higher rates.

2. *A common market*

The common market of the EC, however, goes further than a customs union for it envisages goods and factors of production moving freely within the Community through the operation of the price system; only in this way can the full benefits of the larger market be realised.

But it was recognised that this takes time to accomplish. Member countries had already developed their own individual taxes, welfare benefits, mono- poly policies, methods of removing balance-of-payments imbalances, full

employment policies and so on. Such differences could distort the working of the price system because they would give some members advantages over others. For example, suppose Britain taxed refrigerators but not binoculars. This would weight the possibilities of trade against Italy (which has a comparative advantage in producing refrigerators) and in favour of Germany (which has a comparative advantage in producing high-grade binoculars).

Alternatively, the comparative advantage of some countries may lie in the expertise of the professional services they can provide. Usually this means that such services have to be taken to where the customer is (e.g. know-how regarding property development). There must therefore be mobility of labour within the market, e.g. for property developers.

Policy therefore was directed towards the gradual introduction of 'harmonisation' measures, so that the Single Market could come into effect in 1993 (see p. 308). Examples of such measures are:

(a) *A Common External Tariff* (CET) by which members impose tariffs on imports from non-member countries at the same rates. However, some countries, particularly the developing countries, are given preferential concessions.
(b) *A common agricultural policy* (CAP) – see below.
(c) *Removing barriers to trade and the movement of persons and capital between countries.*
(d) *Uniform rules on competition.* To prevent the distortion of competition in trade, uniform regulations have been introduced to cover price-fixing, sharing of markets and patent rights.
(e) *A common transport policy.* By regulating such items as freight rates, licences, taxation and working conditions, the EC can seek to ensure that transport undertakings compete on an equal footing. Any hidden advantages enjoyed by one country would distort the free movement of goods within the Community.
(f) *Harmonisation of tax systems.* As has already been shown, some standardisation of taxation is necessary in order to remove any 'hidden' barriers to trade. This applies particularly to indirect taxes. In the EC value added tax (VAT) is the basic form of indirect tax, and it is proposed that eventually all member countries will levy it at the same rates.

No proposals exist for harmonising income taxes, but most countries have adopted the 'imputation' system of corporation tax (see p. 271).
(g) *Exchange rate stability.* As we have seen, countries can adjust the prices of imports and exports by varying the exchange rate. If this were allowed within the EC, it could enable a member to obtain a competitive advantage over others by depreciating its currency. Thus through the European Monetary System (EMS) countries agreed to maintain their currency at a fixed exchange rate within narrow limits within the ERM and eventually monetary integration (see pp. 309–10).
(h) *A common regional policy.* Just as one nation cannot allow depressed areas to persist, so the EC is expected to help regions of high unemployment. But the needs of special problem areas, e.g. south Wales and the north-east coast of

the UK, have also to be recognised for depression partly stems from agreed EC policies, such as the reduction of the use of coal for environmental reasons and the contraction of steel-producing capacity.

(i) *A social policy.* This is concerned mainly with securing some uniformity of employment and working conditions embodied in the 'Social Chapter' (see p. 309). Assistance towards retraining is given from the Social Fund.

25.4   Advantages for the UK of belonging to the EC

Several advantages can accrue to countries by forming a common market.

First, it increases the possibility of specialisation. The EC provides a market of 340 million people, larger than that of the USA. This allows economies of scale to be achieved, especially as regards products requiring high initial research expenditure, e.g. computers, drugs, nuclear reactors, supersonic aircraft and modern weapons. These economies should enable EC firms to compete more effectively in world markets.

Second, keener competition in the larger market can result in greater efficiency. Within the EC there are no trade barriers which in effect protect inefficient firms. Free trade means that goods and services can compete freely in all parts of the market and that factors of production can move to their most effi-cient uses, not merely within but also between countries. On the other hand, it must be recognised that protective duties may reduce competition from outside the market.

Third, a faster rate of growth should be achieved as a result of increased econ-omies of scale and competition enjoyed by the EC countries. But it is also possible that the EC generates growth by increasing the *prospects* of growth.

Fourth, there could be significant political benefits. As already explained, the ultimate objective of the original advocates of European cooperation was some form of political union. A Western Europe which could speak with one voice would carry weight when dealing with other major powers, particularly the USA. Moreover, the integration of defence forces and strategy would give its members far greater security. Such benefits, it is held, more than compensate for any loss of political sovereignty (see p. 282).

Fifth, because she is a member of the EC and has a stable political background, the UK can attract investment from countries outside (particularly the USA and Japan) who are anxious to obtain the advantages of producing within the EC.

Sixth, the dynamic growth of the EC enables assistance to be given to its poorer regions and to the less developed countries of the world. Already the UK has been a major beneficiary from the Regional Development Fund, the Social Fund and the Agricultural Guidance and Guarantee Fund.

25.5   Problems facing the UK as a member of the EC

While Britain's membership of the EC can secure important benefits and allow her to influence its future development, it does pose special problems.

1. *The CET could lead to the diversion of trade towards less efficient EC suppliers*

The duties imposed by the customs union may allow firms within the common market to compete in price with more efficient firms outside.

Suppose, for instance, that the same machine can be produced by both the USA and Germany but, because the American firm is more efficient, its machine is 10 per cent cheaper. In these circumstances, Britain would, other things being equal, import from the USA. As a member of EC, however, Britain would have to discriminate against the American machine by the appropriate CET, say 20 per cent. This would make the German machine cheaper, and so trade would be diverted to the less efficient producer.

2. *The CAP is a drain on the community funds*

Before joining the EC Britain imported food at the lowest world market price. In so far as the UK farmer could not make an adequate living by selling at free market prices, British policy consisted of granting *deficiency payments* (financed out of taxation) sufficient to raise the price received by the farmer to a level set out in an Annual Review. The consumer paid a low price for food and the world had free access to the UK market.

But because the Community could not function satisfactorily if the cost of food to consumers differed appreciably in various parts of it, there has to be some equalisation of prices. Yet if this occurred through competition, the market clearing price could destroy many small farmers, particularly in France and Germany. Furthermore, because demand for agricultural products tends to be price-inelastic, even in the short run changes in the conditions of supply, e.g. through a good harvest, can have a damaging effect on farmers' incomes. Even in the long run farmers face relative falling prices since demand for foodstuffs is income-inelastic while supply conditions improve over time through technical innovation.

The CAP supports farmers' incomes by: (a) direct payments per hectare for land sown for cereals, rape and linseed, and also for land compulsorily 'set aside'; (b) subsidies on ewes and beef cattle per head; (c) a guaranteed ('intervention') price for cereals, beef and milk; (d) restricting imports by protective duties at the Community's external frontier.

Three prices are fixed for each product:
 (i) a *target price,* which, it is estimated, will give farmers an adequate return in a normal year;
 (ii) the *intervention price*, at which produce of a specified standard will be bought by the various agencies to prevent the price falling more than 8 per cent below the target price. Thus a farmer can choose between selling his corn on the market or 'putting it into intervention'.
(iii) a *threshold price,* which is the price set for calculating duties on imports when the world price is 10 per cent below the target price.

In practice, giving farmers a guaranteed price above the market clearing price simply encourages overproduction, and stocks accumulate. If these stocks were actually drawn on when harvests were poor, costs would be recouped. In practice, however, improved techniques have so increased supply that it has regularly

exceeded demand. Thus surpluses have accumulated, e.g. butter, beef and corn mountains and milk and wine lakes.

Not only does the system represent an inefficient use of resources, but it is inequitable in that it gives more support to the larger farmer.

3. *'Dumping' surplus produce on world markets injures the less-developed countries and antagonises the USA, Australia, etc.*

Not only have many world producers lost important markets in Europe, but EC surpluses are dumped on world markets. This depresses the price received by all exporting countries, including even the less-developed.

The USA, supported by Canada, Australia and New Zealand, has linked EC agricultural protection with the current Uruguay Round of GATT, making the reduction of tariffs on manufactured goods conditional on an EC reduction in price support for her agricultural produce. While this is in harmony with the British view (see below), it is difficult where the farming lobby is politically important.

A considerable reduction in milk output was achieved by the introduction of production quotas in 1984. As regards cereals, progressive reductions in the real prices received by farmers did not reduce overproduction. Hence in 1993 compulsory 'set aside' was introduced to take 15 per cent of the land under cereals out of production. In addition, intervention prices for cereals are cut by a third, with farmers given income compensation, reducing over three years. Subsidies are also paid on other agricultural products.

4. *There is insufficient control over the community budget*

A Community budget is necessary to meet the costs of administration and policies requiring expenditure, e.g. CAP and regional assistance. There are four main sources – agricultural levies on imported produce, import duties on non-Community goods, a VAT rate up to 1.4 per cent, and payments related to each member's GNP (1.2 per cent in 1992).

Today, over 70 per cent of the budget is spent on agricultural support, and this is increasing. As a result, annual Community expenditure continually exceeds receipts so that the budget revenue has to be increased.

Since the CAP confers greater benefits on countries where agriculture is important (e.g. France), countries depending on industry, e.g. Germany and the UK, lose by having lower receipts from their production of foodstuffs in addition to their budget contribution.

The UK has argued that her net contribution to the EC exceeds that of richer members. This arises because, instead of buying her agricultural imports at world prices, she has to impose a levy if they come from non-EC countries. But this levy goes to the EC budget and not the UK's, with the result that the UK is a net contributor to EC funds.

Her policy, therefore, has been to: (i) negotiate a rebate and seek higher payments from the Regional Fund; (ii) reform the CAP; and (iii) resist moves to increase the budget revenue until extravagance and fraud have been eliminated.

5. *The UK could face an unacceptable loss of political sovereignty* (see below).

25.6   The Single Market: 1995

(a) Measures essential to creating a single market

Although tariffs between members have been scrapped, non-tariff barriers may still obstruct free and fair trade, e.g. transport controls; restrictions on public purchasing, on financial services offered by banks, insurance companies, etc. and on capital movements.

Many of these non-tariff barriers can be overcome by harmonisation of national requirements, e.g. as regards transport safety rules, product standards, acceptable professional standards. Alternatively, the Commission can persuade states to remove national restrictions, e.g. on public purchasing and capital movements. In the last resort major areas of the programme can be implemented by qualified majority voting, although matters of vital national interest, such as taxation, require unanimity.

As far as possible the Commission seeks to eliminate controls by deregulation generally rather than by dealing with each control individually. For instance, if one state imposed an excise duty on a good which was higher than that of another state the market would automatically transfer trade in that good from the dearer country to the cheaper, so that no regulation on duties would be necessary, e.g. cars, where there is now little to be gained by the British motorist buying on the continent and importing.

(b) Economic benefits of the single market

Removing restrictions on trade will allow the full benefits of free trade to be achieved within a market of 340 million people, one of the biggest and richest in the world. More than that, countries outside will also benefit from the increased economic growth of the market since it will make exporting easier for them.

A report prepared by economists for the Commission in 1985 considered that, as a result of:

(a) the removal of frontier barriers affecting trade;
(b) the removal of barriers affecting overall production, e.g. through different standards;
(c) economies of scale; and
(d) intensified competition, the real GNP of the Community should increase by 5 per cent over a six- to eight-year period. Furthermore, as activity responded to the removal of production restrictions, 2 million new jobs could be created. Other employment could result through investment by firms outside the Community seeking to win for themselves the advantages of servicing this large market.

   Consumers, too would benefit from lower prices and more choice especially as regards electrical goods, motor vehicles, pharmaceuticals, food and drink, airlines and financial services.

(c) Conclusions

While the Single Market offers increased opportunities, the tougher competition presents challenges. Large-scale, low-cost producers will benefit at the expense

of smaller higher-cost producers, some of whom could be taken over or go out of business. Even the larger British firms, with their greater reliance for expansion on equity finance through the capital market, may be more vulnerable to take-over or merger.

Certain regions, too, may find the readjustment painful, and the EC must provide extra finance through its Regional, Social and Agricultural Guidance Funds to promote the development of the poorer regions.

The importance of the Single Market to the UK cannot be overemphasised. In value 57 per cent of the UK's exports go to the EC and 52 per cent of her imports come from there.

25.7 Recent developments

(a) The UK and the Maastricht Treaty

At Maastricht in December 1991, the EC countries agreed on a timetable for achieving European Monetary Union (EMU).

An effective unified Single Market can be achieved only if the factors affecting the stability of exchange rates, especially inflation, the balance of payments and the rate of interest, are harmonised by all members. Hence in Stage I all members would join the Exchange Rate Mechanism and maintain the declared exchange rate for their currency within the agreed bands. Controls on capital movements between countries would also be abolished.

Stage 2 would be a transition period, of about three years starting in 1994, when convergence on inflation, interest rates, balance of payments position and fiscal policy would allow all members to achieve stable exchange rates within the narrow band.

Once this had been achieved, declared exchange rates would be locked irretrievably in an exchange rate union (Stage 3). But even this was seen as a transition step for, once fixed exchange rates had been achieved, the cost of exchanging currencies could be eliminated by the EC having a single currency. The Ecu would now become the EC's medium of exchange, not merely a unit of account as at present. It was envisaged that this would be achieved by the end of the century.

The UK obtained the right to opt out of the single currency. The difficulty is that there would have to be an independent European Central Bank (ECB) to take the monetary measures necessary to ensure that the Ecu retained its value. But, because the size of the PSBR influences a country's rate of interest (see p. 239), monetary control would have to extend to fiscal policy, including taxation. Thus UK macroeconomic policy and Parliament's traditional control through holding the purse strings would be largely surrendered to a EC bureaucratic institution supported, if necessary, by majority voting of the member countries.

Such a loss of national sovereignty proved unacceptable to the UK. Furthermore, she did not accede to the social chapter, considering that conditions of employment should be determined in the labour market rather than by general EC rules imposed arbitrarily.

(b) The Exchange Rate Mechanism (ERM)

Stability of exchange rates was to be achieved through the ERM which the UK joined in October 1990. A rate of £1 = 2.95 DM was to be kept within a 6 per cent band either way (as opposed to other members' 2 1/4 per cent band).

But speculation against the £ in September 1992 forced her to suspend membership, leaving the exchange value of sterling to be determined on the foreign exchange market. Italy, Portugal and Spain were also forced to devalue.

Holding the 2.95 DM exchange rate had been achieved by maintaining a relatively high rate of interest. Abandoning the ERM enabled the rate of interest to be reduced to 6 per cent by March 1993, while the devaluation of the £ (2.55 DM, August 1993) made UK exports more competitive. Both stimulated UK recovery from the recession.

However, the persistent refusal of the Bundesbank to lower the German rate of interest because of the economic difficulties created by the re-unification of Germany, meant that in August 1993 other currencies, and in particular the French franc, were unable to keep within the 2 1/4 per cent band. The existence of the ERM was preserved by extending a 15 per cent band to all remaining ten members. Since this in practice simply allows currencies to float freely, the objective of stable currency exchange rates has been postponed indefinitely, and there is now little chance of fulfilling the proposed Maastricht timetable.
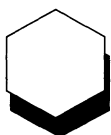
Part IX
Current problems

㉖ Current economic problems and policies of the UK

26.1 Introduction

The economic problems facing the UK at the present time illustrate the point emphasised earlier: because stabilisation policies exhibit a degree of incompatibility, the government has frequently to adjust its priorities according to changes in the economic background.

Moreover, even within a government there may be divergent views as to priorities and policy measures from different wings of the party, e.g. as regards the levels of taxation and government expenditure, the relationship with the EC, the size of overseas aid. Such political considerations may further restrict the policy options open to the government.

26.2 The economy under the Thatcher government

Mrs. Thatcher's government of 1979 emphasised the *control of inflation* objective by: (a) monetarist policy; (b) reducing the PSBR; (c) powers restricting trade unions; (d) supply-side measures.

(a) Monetarist policy

Monetarist policy embodied a *medium-term financial strategy* (MTFS) which limited increases in the money supply as measured by M_3 (broadly M_0 + sterling deposits of banks). But M_3 proved to be a misleading measure. It did not respond as expected to changes in the rate of interest, largely because other sources could provide credit for spending, e.g. building societies, foreign banks.

In 1987, M_0 – an 'indicator', not a target – was substituted to monitor changes in the money supply. Monetary restraint measures were based on shadowing the £/DM exchange rate. This was formalised in October 1990 when the UK joined the ERM committed to maintaining a £1 = 2.95 DM exchange rate with some flexibility within a 6 per cent band either way.

(b) The PSBR

Reducing the PSBR was necessary because it creates monetary problems through the borrowing necessary to cover it.

If the government borrows by selling Treasury bills, the cheapest method, it runs into difficulties. The major holders of Treasury bills are the commercial banks, who buy them because an increased offering forces up the yield. Such purchases, however, directly add to the banks' liquid assets, allowing them to increase their deposits. In other words, short-term bank borrowing involves an inflationary increase in the money supply.

As a result, the government has to rely on long-term borrowing, selling medium- and long-term bonds in the market to the *non-bank* sector, the institutions and private purchasers. Since such sources of funds rely mainly on current saving, this method is not inflationary. The difficulty, however, is that extra bonds can only be disposed of at a lower price – that is, by a rise in the long-term rate of interest. This has the overall effect of discouraging investment, thereby retarding growth. Nor is this all. Interest payments on this borrowing add to the PSBR. Furthermore, inasmuch as higher interest rates attract funds from abroad, the money supply is increased.

It is important to note the real significance of this. Whereas budgetary policy was the main weapon of Keynesian demand management, it is now merely a support to monetary policy, enabling liquidity to be controlled at a lower rate of interest than would otherwise be necessary.

By exercising a stricter control over government spending, the Thatcher government succeeded in reducing the PSBR. It was also helped by the 'privatisation' proceeds of selling government assets in the nationalised industries. With an expanding economy, taxation yields increased and employment-related payments fell. During 1987–9 the government achieved an annual surplus of revenue over expenditure.

(c) Wage restraint

To support its anti-inflation objective, the government curbed trade union legal powers and limited public sector pay increases. Moreover a 13.5 per cent rate of interest led in 1986 to 3.25 mn unemployed, and this further weakened labour's bargaining position.

On the other hand, it is doubtful whether its supply-side measures acted quickly enough or were large enough to have any significant impact on 'stagflation'.

(d) Supply-side measures – see pp. 250–1

26.3   Recession 1989–93

The turning point in the expanding economy can be traced back to the Chancellor of the Exchequer's apprehensive response to the crash in the world's stock markets in October 1987. An easier monetary and fiscal policy proved a blunder. Imprudent bank lending financed a property boom, and a fall in the personal saving–income ratio from 12.2 to 4.4 per cent fuelled a rapid rise in house prices and an increase in the inflation rate to nearly 8 per cent in 1989. Consequently the rate of interest was increased, rising to 15 per cent in October 1989.

The high cost of borrowing led to a *cumulative recession* through reduced investment, a slump in the demand for houses, a rise in the personal saving ratio, the failure of businesses (particularly of over-geared property companies) and bank write-offs of losses.

In addition, by joining the ERM in October 1990, the UK had to retain a relatively high rate of interest to support an overvalued £1 = 2.95 DM exchange rate.

Thus neither a consumer-led nor an export-led recovery materialised quickly, especially as the recession spread worldwide.

It is against this background that we discuss briefly the present-day problems of the UK. To simplify, they are considered under four separate headings: inflation, the PSBR, balance of payments difficulties and unemployment. But since there are links one with the other, they have to be combined in an overall government economic strategy.

26.4   The main economic problems

(a) Inflation

Today (June 1993) the rate of inflation is the lowest for thirty years – the headline rate being 1.2 per cent and the underlying rate (which excludes mortgage payments) being 2.8 per cent. Suggested reasons for such a low rate are:

   (i)   unemployment of 2.9 million has curbed demands for wage increases;
   (ii) aggregate demand remains low, consumers having increased their saving ratio to around 11 per cent and firms (particularly those concerned with property development) being reluctant to invest. In fact, increases in retail sales and house purchase are now regarded as welcome 'green shoots' of economic recovery.

Current government policy is to keep the underlying rate of inflation within a 1 to 4 per cent range, with the long-term aim of 2 per cent or less. This is to be achieved largely by monetary measures based on:

   (i)   a *target* annual growth range for M_0 of 0 to 4 per cent;
   (ii) a *monitored* range for M_4 of 4 to 8 per cent;
   (iii) other indicators, e.g. house prices, exchange rate movements;
   (iv) interest rate adjustments to depend on prospective inflationary trends, with the Bank of England publishing quarterly reports on its views.

Public sector pay increases in 1993 were restricted to $1\frac{1}{2}$ per cent. To keep within the stated limits the main weapon will be the rate of interest. This means that the determination of the government to make the control of inflation its first objective could be sorely tested by pressure from industry to reduce the rate in order to speed up the recovery from recession.

(b) The PSBR

For 1993–4 the government has forecast a PSBR of £45 bn, equal to 9 per cent of GDP. As we have seen, this has to be covered by borrowing. If such borrowing is from the *banking* sector, it adds to inflationary pressure; if it is from the non-bank sector, it is likely to mean a rise in the long-term rate of interest. Thus, even with inflation under control, the large PSBR reduces the ability of the government to lower interest rates.

But getting down the PSBR by increased taxation and reduced government expenditure comes up against both economic and political difficulties.

Increasing direct taxation could cut off the 'green shoots' of economic recovery by reducing aggregate demand and stifling incentives. Nor does it accord at present with the political philosophy of the Conservative Party. Increasing

indirect taxation adds to inflation risks by raising the cost of living, and is regressive. Since, in addition, many rank and file Conservative supporters would be adversely affected, and vote accordingly (as with imposing VAT on domestic heating), any future indirect tax increases could be largely confined to the four basic sources: tobacco, alcohol, motoring and gambling.

To some extent the PSBR is 'cyclical' in that, as the economy recovers, so will the proceeds of taxation increase and expenditure on unemployment-related benefits fall. But much expenditure has to be regarded as 'structural' in that it follows from government commitments – social welfare payments, the health service, education, law and order, etc. Even though reduced defence spending has been made possible by the ending of the Cold War (the 'peace dividend'), additional commitments have arisen, in financing UN peace-keeping operations, the Russian transition to a market economy, and the rehabilitation of refugees.

The government, therefore, is carrying out a review of all main headings of expenditure and suggesting the various policy options. One question which does have to be answered is whether there should be some reform of the social security system. Spending on state welfare benefit payments, e.g. retirement pensions, invalidity benefits, income support and housing allowances, has increased nearly seven-fold in real terms since the welfare state was set up and now accounts for almost one-third of all public spending. Even assuming an annual 2.5 per cent real growth in GDP and unemployment down by a quarter, the built-in bill will rise by an average of 2.6 per cent a year from £74.1 bn in 1992–3 to £88 bn by the end of the century. In short, spending on welfare benefits is outpacing growth in the economy as a whole, highlighting the difficult task which the government faces in reducing the £45 bn PSBR.

(c) Balance of payments difficulties

Throughout the twentieth century the UK has been losing her share of world trade as other countries have industrialised and competed in world markets. Moreover, with overseas investments being sold to help pay for two World Wars, income from invisibles dwindled.

For a time, North Sea oil transformed the situation. By 1979 the UK was self-sufficient in oil and, as production expanded, so she became a net exporter. Surpluses on the balance of payments 1986–9 were used to accumulate overseas assets, which rose from £5 bn in 1979 to £160 bn by 1989.

It can be argued, however, that such overseas investment would have been better spent in re-equipping British industry so that she could compete more effectively in price and quality with those countries, such as Japan and Germany, whose industries were technically more advanced.

That this did not happen largely resulted from the effect of the strength of the UK oil exports on the foreign exchange market. The £ sterling became virtually a petro-currency, with its exchange rate largely dependent on the price of oil. Thus when the price of oil trebled in 1984 to $30 a barrel, the £ appreciated to $2.50. This reduced still further the competitiveness of British goods, and manufacturing industries in particular were forced to cut production. By 1983, trade in manufactured goods had moved from surplus to deficit.

Since 1985 the value of the UK's oil exports has been falling, partly through our reduced production and partly through a drop in the world price of oil to $14 a barrel. This means that, to correct today's (1992) balance of payments deficit of £8.6 bn the value of other exports has to be increased (see below) or imports decreased, chiefly by reducing foreign penetration of the home market.

The 15 per cent depreciation of the £ sterling when the UK left the ERM in September 1992 gave an immediate price advantage to her exports. But depreciation can only be successful in the long term if wage-costs per unit can be reduced relative to those of her competitors, and if the resulting increased overseas demand for her goods can be supplied. The first requires increased productivity and restraint in wage rises – even though higher prices on the home market are a likely result of the higher cost of imports. The second is possible since there is slack in the economy resulting from the recession.

(d) Unemployment

UK unemployment has increased from 1.8 mn in 1989 to 2.9 mn in June 1993, 10.4 per cent of the working population.

Some of this unemployment is cyclical, the result of recession. But even with recovery, it is hardly likely to fall by more than a quarter by the end of the century, leaving over two million workers unemployed, many of whom would be classified as 'long-term' in that they have not worked for over a year. Welfare benefits do not compensate for the hopelessness experienced by such workers, especially those over fifty years of age who have been made redundant as financial service industries have contracted, machines have replaced labour, and previously important industries such as coal, ship-building and steel, have contracted.

Dealing with this structural unemployment was discussed in Chapter 21 – introducing new industries to districts where employment is particularly high (e.g. through Enterprise Zones), financial assistance in establishing self-employment businesses. All are likely to require training of young workers and the unemployed in new skills.

As regards cyclical unemployment, because of the serious balance of payments deficit, any consumer-led recovery must be quickly followed by export-led growth. This requires : (i) a general recovery from the world recession, especially as regards the EC countries and the USA; (ii) if conditions permit, a further reduction in the rate of interest; (iii) improved export competitiveness; (iv) a restructuring of British industry.

Improved competitiveness would come from lower relative wage costs per unit of output. The outlook is promising. Industry has eliminated surplus capacity, and improved labour relations have diminished trade union power and enabled firms to concentrate on long-term production strategy rather than on settling disputes. A consistent government economic policy which checks inflation and holds down interest rates should encourage investment in new products and processes.

These factors should also lead to more foreign investment in the UK, especially as firms (Japanese, for instance) seek to establish production bases

within the EC. Indeed their improved management and production techniques could 'spin-off' onto established domestic firms.

British production must be restructured towards manufacturing industries, for these have a possible tradeable export content of 100 per cent (compared with services' 20 per cent), thereby reversing the decline of the last 20 years as importers have penetrated the home market. In particular, the UK should concentrate on developing and exporting new products having a high income-elasticity of demand. In addition UK firms must endeavour to regain the domestic market.

Index